Probabilistic Engineering Mechanics 63 (2021) 103115

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

journal homepage: www.elsevier.com/locate/probengmech

# Are seismic fragility curves fragile?

M. Grigoriu [a,*], A. Radu [b]

[a] Cornell University, Ithaca NY 14853–3501, USA
[b] Global Parametrics, London, UK

ARTICLE INFO

*Keywords:*
Conditional probability
Conditional response spectrum
Fragility curves
Fragility surfaces
Ground acceleration
Stochastic process

ABSTRACT

A hypothetical seismic site is constructed for which the probability law of the seismic ground acceleration process $X(t)$ is specified. Since the seismic hazard is known, the performance of the incremental dynamic analysis- (IDA) and multiple stripe analysis- (MSA) based fragilities, which are used extensively in Earthquake Engineering, can be assessed without ambiguity. It is shown that the IDA- and MSA-based fragilities are unsatisfactory for moderate and large seismic events, are sensitive to the particular parameters used for their construction, and may or may not improve with the sample size. Also, the usefulness of the optimization algorithms for selecting ground motions records is questionable.

## 1. Introduction

Fragilities are conditional probabilities $P(\mathrm{Y} > y_{\mathrm{cr}}|\text{seismic hazard})$ that engineering demand parameters $Y$ exceed critical levels $y_{\mathrm{cr}}$ for given site seismic hazard. The site seismic hazard is specified in this study by stochastic processes whose samples are viewed as ground acceleration time histories, with rates of occurrence defined by a hypothetical system of linear faults with specified parameters. It is common to characterize the site seismic hazard by intensity measures (IMs), e.g., peak ground accelerations or response spectral accelerations, so that the fragilities $P(\mathrm{Y} > y_{\mathrm{cr}}|\text{seismic hazard})$ are approximated by the conditional probabilities $P(\mathrm{Y} > y_{\mathrm{cr}}|\mathrm{IM})$, referred to here as ***current fragilities***. Simplicity is the main feature of IMs. Crude characterization of seismic hazard is a notable limitation of IMs which, as shown in this study, is likely to affect the usefulness of crude fragilities.

It is consensus of the Earthquake Engineering community that IMs have to be such that fragilities conditional on these measures approximate satisfactorily actual fragilities, i.e., $P(\mathrm{Y} > y_{\mathrm{cr}}|\mathrm{IM}) \simeq P(\mathrm{Y} > y_{\mathrm{cr}}|\text{seismic hazard})$. Specifically, IMs are required to be *efficient*, i.e., the conditional random variable Y|IM has small variance, *sufficient*, i.e., the conditional random variables Y|IM and Y|(seismic hazard) have similar distributions, and *scale robust*, i.e., structural responses obtained by scaling records are unbiased [1–3].

IMs satisfy approximately some but not all of the above requirements [1–3]. Moreover, their predictive power decreases with the size of seismic events [3–5]. These are significant limitations which suggest that current fragilities $P(\mathrm{Y} > y_{\mathrm{cr}}|\mathrm{IM})$ may provide unsatisfactory approximations of actual fragilities $P(\mathrm{Y} > y_{\mathrm{cr}}|\text{seismic hazard})$ and that their quality deteriorates with the size of seismic events. The quality of current fragilities $P(\mathrm{Y} > y_{\mathrm{cr}}|\mathrm{IM})$ is further affected by the use of small sets of scaled ground acceleration records which are selected via optimization based on questionable objective functions [6].

Intuition suggests that current IMs, e.g., response spectra, can only provide a crude characterization of the seismic hazard since maxima of responses of simple linear oscillators and of complex nonlinear structural systems to the same ground motions may differ significantly. This intuitive observation is confirmed in [4] by showing that maxima of responses of linear oscillators and even simple nonlinear dynamical systems are weakly dependent.

Our main objective is to assess the performance of the current methods for fragility analysis. A postulated site and seismic hazard is used to achieve these objectives. We recognize that this setting is a simplified version of reality. Yet, it has the advantage that truth is known so that potential pitfalls of current methods for fragility analysis can be identified without ambiguity. If the current fragilities are unsatisfactory in this setting, their usefulness is questionable. We also note that the postulated hazard (1) is consistent with physics, in the sense that the frequency content of the seismic ground acceleration depends on the moment magnitude and the source-to-site distance, and (2) is more realistic than seismic hazards considered elsewhere, e.g., the seismic hazard in [7] depends only on the moment magnitude.

The performance of current fragilities has been investigated extensively by using actual and synthetic ground motions. The synthetic ground motions are samples of postulated stochastic processes [7,8] and the seismic hazard is characterized by IMs, although IMs and engineering design parameters are weakly dependent when dealing with large seismic events [3,4,8]. Accordingly, current fragilities, which are probabilities of engineering design parameters exceeding limit levels conditional on IMs, are not informative for large seismic events.

* Corresponding author.
*E-mail address:* mdg12@cornell.edu (M. Grigoriu).

https://doi.org/10.1016/j.probengmech.2020.103115
Received 26 August 2020; Received in revised form 24 October 2020; Accepted 19 November 2020
Available online 25 November 2020

In contrast, this study defines fragilities as conditional probabilities of engineering design parameters exceeding limit levels conditional on the defining parameters of the ground motion stochastic process, e.g., earthquake magnitude and site-to-source distance for the model in [9] under specified soil conditions.

It is shown that IDA- and MSA-fragilities provide limited information on structural performance for moderate and large seismic events, are sensitive to implementation procedures, seem to converge/diverge with the sample size depending on the value of the critical threshold $y_{\rm cr}$, and are based on IMs which are crude descriptors of seismic hazard. Also, the usefulness of the algorithms for selecting ground motions is questionable since fragilities based on samples selected by these algorithms and on randomly selected ground motions are similar. Fragility surfaces, i.e., probabilities of events $\{Y > y_{\rm cr}\}$ conditional on the defining parameters of the ground acceleration process, e.g., moment magnitude $M$, site-to-source distance $R$, and soil properties, are the only realistic metrics of structural performance.

The following sections define the site seismic hazard (Section 2), illustrate limitations of IMs (Section 3), develop reference fragility curves/surfaces and construct current IDA- and MSA-fragility curves (Section 4), assess the performance of IDA- and MSA-fragilities (Section 5), and present concluding remarks (Section 6).

## 2. Seismic hazard

Seismic events at a site are generated by ruptures along seismic faults in its proximity which occur at random times and random locations along these faults. We consider a hypothetical seismic site, develop a probabilistic model for site ground acceleration process, and construct a set of site ground acceleration time histories which define the site seismic hazard, i.e., the setting in Earthquake Engineering practice. The hypothetical seismic scenario herein has one essential feature to perform a critical review on seismic fragility, that all its parameters, including the probability laws that govern the occurrence and the time history processes of ground motions are known. This utopian scenario may not follow all the features of realistic seismic hazard scenarios, but the development of a perfect hazard model is beyond the scope of the paper. This simplified hazard definition is not a limitation, since our main point is that a robust method for fragility development which works in a realistic complex situation should also work in the simplistic scenario of this paper.

### 2.1. Seismic activity matrix (SAM)

Consider the site in the left panel of Fig. 1 which is affected by three seismic faults. It is assumed that (1) the seismic events occur at random times defined by a homogeneous Poisson process of intensity $\lambda = 1$ event/year, (2) the seismic events are caused by ruptures along the faults 1, 2, and 3 which occur with probabilities $q_1$, $q_2$, and $q_3$, (3) the seismic point sources are randomly located along the faults, and (4) the magnitudes $M$ of seismic events are independent samples of the Gutenberg–Richter law described by the density [10]

$$f_{M,k}(m) = \frac{b_k \ln(10)\, 10^{-b_k (m-m_{\min,k})}}{1 - 10^{-b_k (m_{\max,k}-m_{\min,k})}}, \tag{1}$$

where $m_{\min,k}$ is the minimum magnitude of interest, set to 5 for all faults, and $m_{\max,k}$, $k = 1,2,3$, is an upper bound of the magnitude for each fault, and $b_k$ is a parameter which controls the right tail of the Gutenberg–Richter distribution. Numerical values of the parameters in Eq. (1) are in Table 1. The entries $(x_{1,k}, y_{1,k})$ in the table are the coordinates of the end points of the three faults.

The parameters associated to the three fault lines that define the hazard scenario are customary and may be not respect the relationship between fault size and moment magnitude but it is considered adequate for our objective. The right panel of Fig. 1 shows a two-dimensional histogram of moment magnitude $M$ and site-to-source distance $R$, referred to as seismic activity matrix (SAM). The histogram is based on 10,000 seismic events. It provides information on the likelihood of observing seismic events corresponding to various $(M, R)$-values. Such statistics are common in Earthquake Engineering, and identical or similar information about realistic hazard scenarios in the United States, for example, can be obtained from sources such as the United States Geological Survey (USGS), using their earthquake-probability [11] or uniform-hazard tools [12].

**Table 1**
Fault Properties.

| Fault $k$ | $q_k$ | $m_{min,k}$ | $m_{max,k}$ | $b_k$ | $(x_{1,k}, y_{1,k})$ | $(x_{2,k}, y_{2,k})$ |
|---|---|---|---|---|---|---|
| 1 | 0.2 | 5 | 7 | 1.10 | $(-20, -40)$ | $(120, -15)$ |
| 2 | 0.5 | 5 | 8 | 0.90 | $(-80, -20)$ | $(20, -25)$ |
| 3 | 0.3 | 5 | 8 | 0.85 | $(-20, 50)$ | $(150, 60)$ |

### 2.2. Seismic ground acceleration process

Denote by $(m, r)_j$ the centers of the cells of the SAM in Fig. 1 and by $p_j$ the probability that, given the occurrence of a seismic event, its magnitude $M$ and distance $R$ belong to cell $j = 1, \ldots, N$, where $N = 300$, $M \in [5.0, 8.0]$, and $R \in [0, 200]$ km.

To complete the probabilistic model of the site seismic hazard, we now specify the probability law of the site seismic ground acceleration process. The model for this process relies on the specific barrier model (SBM) [13,14], which is a seismological model that provides the frequency content of ground motions as functions of $(M, R)$, in the form of a power-spectral density function $g(\nu; M, R)$, given some other fixed conditions, such as the shear-wave velocity or the seismological regime. The SBM was calibrated to regional data in [15], and was further improved to be consistent with site-specific records in [16,17]. Of course, other substitute stochastic models such as the ones proposed by Rezaeian and Kiureghian [18], Tsioulou et al. [19], or Vlachos et al. [20], or more complex physics-based models such as Goda et al. [21] can be used to simulate earthquakes as functions of $(M, R)$.

The construction of this ground-motion process in our study uses developments in [9], and is based on the following assumptions.

(1) The active $(m, r)_j$ cell during an arbitrary seismic event is a sample of the multinomial random variable $(V_1, \ldots, V_N)$ with parameters $(p_1, \ldots, p_N)$, i.e., the outcome of the experiment of rolling an $N$-sided die with side probabilities $(p_1, \ldots, p_N)$ and
(2) The site seismic ground acceleration processes $X_j(t)$, $j = 1, \ldots, N$, corresponding to the cells of the activity matrix are independent Gaussian processes whose laws are given by the seismological model in [15]. Hence, the site ground acceleration during a seismic event is the stochastic process

$$X(t) = \begin{cases} X_1(t), & \text{with probability } p_1 \\ \quad\vdots & \\ X_j(t), & \text{with probability } p_j \\ \quad\vdots & \\ X_N(t), & \text{with probability } p_N \end{cases} \tag{2}$$

Various probabilistic models can be used for the processes $X_j(t)$ [16]. We use the SBM model of [15] since it is based on seismological arguments and can be updated if additional site ground-motion records, regarded as samples of $X(t)$, are available [17]. According to this model, $X_j(t)$ are non-stationary Gaussian processes defined by

$$X_j(t) = h_j(t)\, Z_j(t), \quad j = 1, \ldots, N, \tag{3}$$

where $h_j(t)$ is a deterministic modulation function with support $[0, \tau_j]$ and $Z_j(t)$ is a zero mean stationary Gaussian process with one-sided spectral densities $g_j(\nu)$, $\nu \geq 0$, which depends on the site-to-source

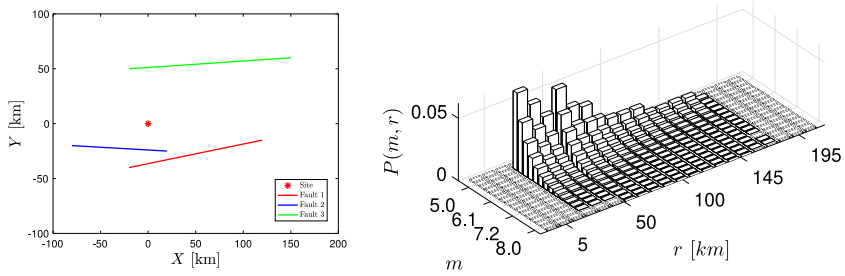


**Fig. 1.** Hypothetical site and seismic activity matrix (SAM) (left and right panels).

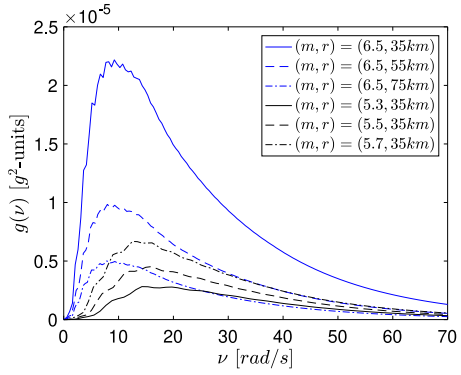


**Fig. 2.** Comparison of SBM power spectral densities for different pairs of $(m, r)$.

distances $r_j$, source magnitudes $m_j$, and other parameters, e.g., soil conditions.

Fig. 2 shows spectral densities of the processes $Z_j(t)$ for a site with type-$D$ NEHRP (National Earthquake Hazards Reduction Program) soil characterized by an average shear velocity in the top 30 m of soil of $v_{s30} = 310$ m/s. The frequency content and ordinates of the spectra vary significantly with $(M, R)$. The modulation function has the form

$$h_j(t) = a\, t^b \exp(-c\, t) \tag{4}$$

where the scalar coefficients $a$, $b$, and $c$ are also outputs of the SBM and depend on $(M, R)$ [15].

According to the above models, site seismicity is completely defined by the SAM in Fig. 1, the Gutenberg–Richter law of Eq. (1), and nonstationary Gaussian process $X(t)$ in Eqs. (2)–(4). For given soil conditions, the site seismicity is specified by one member of the family of processes $X_j(t)$, $j = 1, \ldots, N$, whose selection depends on $(M, R)$. This means that fragilities can be defined precisely as probabilities of the events $\{Y > y_{cr}\}$ conditional on $(M, R)$ for given soil conditions.

### 2.3. Reference ground acceleration records

The following two-step Monte Carlo algorithm has been implemented to generate $n = 10{,}000$ independent seismic ground acceleration records for the site in Fig. 1.

- *Step 1:* Generate $n$ independent samples of the multinomial variable $(V_1, \ldots, V_N)$ with $N$ states of probabilities $(p_1, \ldots, p_N)$. The distribution of this variable is

$$P(V_1 = v_1, \ldots, V_N = v_N) = \frac{n!}{v_1! \cdots v_N!}\, p_1^{v_1} \cdots p_N^{v_N}, \tag{5}$$

  where $v_j$ are non-negative integers such that $\sum_{j=1}^N v_j = n$. It gives the probability that the cells indexed by $\{1, \ldots, N\}$ are selected $\{v_1, \ldots, v_N\}$ times in $n$ trials. The MATLAB function **mnrnd**$(n, p)$ can be used to generate samples of the distribution in Eq. (5), where $p = (p_1, \ldots, p_N)$. The outcome of this function is the active cell $k_i \in \{1, \ldots, N\}$ at each trial $i = 1, \ldots, n$. For example, if $k_i = 1$, seismic cell 1 is selected at the $i$th trial.
- *Step 2:* Generate $n$ independent samples of $X(t)$, i.e., single independent samples of the processes $X_{k_i}(t)$, $i = 1, \ldots, n$. Standard algorithms can be used to generate samples of the stationary Gaussian processes $Z_{k_i}(t)$ in Eq. (3), e.g., methods based on the spectral representation and the sampling theorems [22].

The resulting set of $n$ independent samples $\{x_1(t), \ldots, x_n(t)\}$ of $X(t)$, referred to as the ***reference set*** of ground accelerations, is used to characterize the site seismic hazard. This set of synthetic records is similar to the set of ground motion time histories recorded during seismic events at a site. They are used in the subsequent sections for fragility analysis.

We note that the reference set of $n = 10{,}000$ synthetic records is large relative to the sets of site ground acceleration records available at most sites. Yet, it is insufficient to characterize accurately the seismic hazard, as illustrated in a subsequent section. The remainder of this section examines potential difficulties related to sample size $n$, i.e., the size of the available set of ground acceleration records.

Let $V = (V_1, \ldots, V_N)'$ denote the $N$-dimensional column random vector of multinomial variables with probability in Eq. (5). The samples of $V$ are column vectors whose components are equal to zero except for one which is equal to one. Let $V^{(i)}$, $i = 1, \ldots, n$, be independent copies of $V$ and define the vector-valued estimator

$$\hat{P} = \frac{1}{n} \sum_{i=1}^n V^{(i)} = \frac{1}{n} \begin{bmatrix} \sum_{i=1}^n V_1^{(i)} \\ \vdots \\ \sum_{i=1}^n V_N^{(i)} \end{bmatrix} \tag{6}$$

of the probabilities $p = (p_1, \ldots, p_N)'$. Note that $N_j := \sum_{i=1}^n V_j^{(i)}/n$ gives the random number of occurrences of source $j = 1, \ldots, N$ in $n$ trials scaled by $n$ so that $E[N_j] = p_j$, $\mathrm{Var}[N_j] = p_j(1 - p_j)$, and $\mathrm{Cov}[N_j, N_r] = -n\, p_j\, (1 - p_r)$ [23] (Chap. 11). The mean and covariance matrices of the estimator $\hat{P}$ are

$$E\big[\hat{P}\big] = \frac{1}{n} \sum_{i=1}^n E\big[V^{(i)}\big] = p \quad \text{and}$$

$$E\big[(\hat{P} - p)\,(\hat{P} - p)'\big]_{st} = \frac{1}{n^2}\, E\left[\sum_{i=1}^n \big(V_s^{(i)} - p_s\big) \sum_{j=1}^n \big(V_t^{(j)} - p_t\big)\right]$$

$$= \frac{1}{n}\left[p_s\,(1 - p_t)\,\delta_{st} - p_s\, p_t\,(1 - \delta_{st})\right], \quad s, t = 1, \ldots, N, \tag{7}$$

where $\delta_{st} = 1$ and 0 for $s = t$ and $s \neq t$. These moments show that $\hat{P}$ is an unbiased and consistent estimator since $E\big[\hat{P}\big] = p$ and $E\big[\sum_{i=1}^n \big(V_s^{(i)} - p_s\big)\, \sum_{j=1}^n \big(V_r^{(j)} - p_r\big)\big]_{st} \to 0$ as $n \to \infty$.

These considerations show that the mean and variance of the random number of ground accelerations associated with cell $j$ in $n$ trials are $n\, p_j$ and $(1/p_j - 1)/n$. For example, we will see on average 10 samples from the cell $j$ of the SAM matrix with $p_j = 0.001$. Since the coefficient of variation of this number is $\mathrm{Cov}[\hat{P}_j] = \sqrt{(1/p_j - 1)/n} = 0.31$, the observed number $\hat{P}_j$ of records from this cell exhibits large sample-to-sample variation.

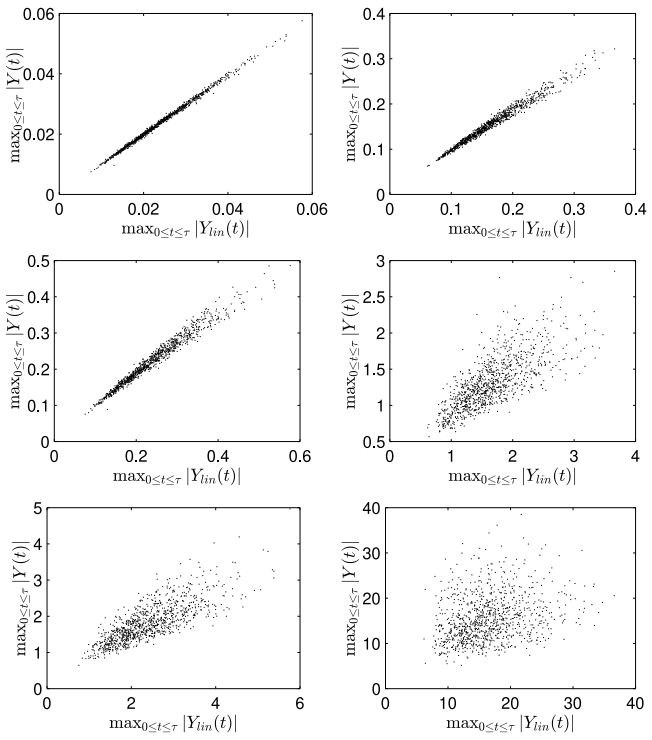


**Fig. 3.** Scatter plots of $\max_{0\le t\le\tau}\{|Y_{\text{lin}}(t)|\}$ and $\max_{0\le t\le\tau}\{|Y(t)|\}$ for $(m = 5.3, r = 35$ km) (left panels) and $(m = 6.5, r = 35$ km) (right panels) for scaling factors $\kappa = 0.1$, 1, and 10 (top, middle, and bottom panels).

## 3. Current intensity measures

Ordinates of response spectral accelerations for specified periods and damping ratios are currently used as intensity measures (IMs) for fragility analysis. As previously stated, intuition suggests that these IMs are unlikely to predict the performance of complex nonlinear dynamical systems. Theoretical arguments based on concepts of the extreme value theory (EVT) [4] and numerical illustrations [3,5] provide quantitative support to this intuition. Recent studies show, in agreement with findings in [4,5], that IMs can yield bias estimates of structural responses [2,24]. This section presents additional numerical examples which further illustrate limitations of IMs.

Let $Y_{\text{lin}}(t)$ and $Y(t)$ denote the displacement of a linear single-degree of freedom system and a scalar response of a multi-degree of freedom nonlinear structural system of arbitrary complexity subjected to the same ground acceleration process $X(t)$. Following practice, the natural period and damping ratio of the linear oscillator are tuned to properties of the nonlinear system, e.g., they match the period and the damping ratio of the first mode of vibration of the nonlinear system for small oscillations [6].

The responses $Y_{\text{lin}}(t)$ and $Y(t)$ are dependent processes as solutions of differential equations to the same input $X(t)$. However, their dependence is weak since (1) $Y_{\text{lin}}(t)$ and $Y(t)$ have different frequency bands as linear systems filtered out frequencies while nonlinear systems, generally, create new frequencies and (2) $Y_{\text{lin}}(t)$ is Gaussian if the input is Gaussian while $Y(t)$ is not irrespective of the properties of $X(t)$ so that extremes of $Y_{\text{lin}}(t)$ and $Y(t)$ are likely to differ significantly. The discrepancy between the probability laws of $Y_{\text{lin}}(t)$ and $Y(t)$ strongly suggests that accurate predictions of responses of nonlinear systems from corresponding responses of linear systems are unlikely.

We further illustrate the discrepancy between the response processes $Y_{\text{lin}}(t)$ and $Y(t)$ for a Bouc–Wen single degree of freedom system whose state $Y(t)$ satisfy the differential equation

$$\dot{Y}(t) + 2\zeta\nu\dot{Y}(t) + \nu^2(\rho Y(t) + (1-\rho)W(t)) = \kappa X(t), \quad t \in [0,\tau],$$
$$\dot{W}(t) = -\gamma\dot{Y}(t) + \alpha|W(t)|^{\eta-1}W(t) + \beta\dot{Y}(t)|W(t)|^{\eta}, \tag{8}$$

where $\zeta$, $\nu$, $\alpha$, $\beta$, $\gamma$, $\rho$, $\eta$ are model parameters, $W(t)$ denotes the hysteresis response of the Bouc–Wen oscillator, $\kappa > 0$ is a scale factor, and $X(t)$ is the seismic ground acceleration process defined by Eq. (3) and is scaled by $\kappa = 0.1$, $\kappa = 1.0$ and $\kappa = 10$. Note that the solution for $\rho = 1$ is that of a linear oscillator with natural frequency $\nu$ and damping ratio $\zeta$, i.e., the process $Y_{\text{lin}}(t)$. The following numerical results are for $\nu = 2\pi$, $\zeta = 0.05$, $\alpha = 0.5$, $\beta = 0.5$, $\gamma = 1$, $\rho = 0.1$, $\eta = 1$, $\kappa = 1$, and $n = 10{,}000$ samples of $X(t)$ corresponding to $(m = 5.3, r = 35$ km) and $(m = 6.5, r = 35$ km). The duration $\tau$ of the seismic input depends on $(M, R)$ and can be found in [9].

Fig. 3 shows scatter plots of response maxima of the linear and the Bouc–Wen systems, i.e., the random variables $\max_{0\le t\le\tau}\{|Y_{\text{lin}}(t)|\}$ and $\max_{0\le t\le\tau}\{|Y(t)|\}$, corresponding to the reference set of $n = 10{,}000$ ground acceleration samples. The left, middle, and right panels are for scaling factors of $\kappa = 0.1$, 1, and 10. The left and right panels are for $(m = 5.3, r = 35$ km) and $(m = 6.5, r = 35$ km). Large responses $\max_{0\le t\le\tau}\{|Y_{\text{lin}}(t)|\}$ of the linear oscillator are satisfactory predictors of the corresponding responses $\max_{0\le t\le\tau}\{|Y(t)|\}$ of the Bouc–Wen system for small seismic events ($\kappa = 0.1$), an expected result since the Bouc–Wen oscillator behaves as a linear oscillator for small inputs. On the other hand, $\max_{0\le t\le\tau}\{|Y_{\text{lin}}(t)|\}$ is a pure predictor of $\max_{0\le t\le\tau}\{|Y(t)|\}$ for moderate and large seismic events ($\kappa = 1$ and 10). This is a significant limitation since damages and potential life losses occur during large seismic events.

The practical implication of the plots of Fig. 3 is significant. The plots show that the ground motions which maximize the responses of linear oscillators may not maximize the responses of nonlinear systems. This means that the selection of representative ground motions based on response maxima of linear systems, i.e., response spectra, is questionable.

## 4. Fragility analysis

Suppose that seismic hazard at a site is described by the reference set $\{x_1(t), \ldots, x_n(t)\}$ of ground acceleration records, i.e., $n$ independent samples of the ground acceleration process $X(t)$ in Eq. (2). We define and construct two types of fragility curves, referred to as ***reference*** and ***current fragility curves***. Both types of fragilities have access to the same information, the reference set $\{x_1(t), \ldots, x_n(t)\}$ of seismic ground acceleration records. We also construct ***reference fragility surfaces*** which use sets of records larger than $\{x_1(t), \ldots, x_n(t)\}$.

### 4.1. Reference fragility curves

These fragilities are derived from the entire set of unscaled reference records $\{x_1(t), \ldots, x_n(t)\}$. The construction of these fragility curves involves the following four steps.

- *Step 1:* Construct the pseudo-acceleration response spectra $\{S_{a,i}(T)\}$, $T \ge 0$, of the unscaled seismic records $\{x_i(t)\}$, $i = 1, \ldots, n$, for a specified damping ratio $\zeta$, e.g., $\zeta = 0.05$.
- *Step 2:* Partition the spectral range $[\min\{S_{a,i}(T_0)\}, \max\{S_{a,i}(T_0)\}]$ at the reference period $T = T_0$ in $\bar{k}$ equal or unequal intervals $I_k$, $k = 1, \ldots, \bar{k}$. The spectral range defines the range of possible values of IMs.
- *Step 3:* Partition the set of ground motions in $\bar{k}$ subsets by using the following rule. The ground motion $x_i(t)$ is assigned to the interval $I_k$ if $S_{a,i}(T_0)$ is in $I_k$. Denote by $n_k$ the number of ground motions assigned to $I_k$.
- *Step 4:* Calculate structural responses $y_i(t)$ to the ground motions $x_i(t)$ of the reference set and estimate fragilities in each interval

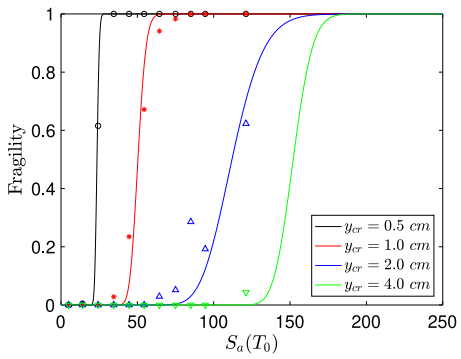


**Fig. 4.** Reference fragility curves calculated for the Bouc-Wen system for $y_{cr} = 0.5$, $y_{cr} = 1.0$, $y_{cr} = 2.0$, and $y_{cr} = 4.0$ cm.

$I_k$ from

$$P_{f,k}(y_{cr}) = \frac{1}{n_k}\sum_{i=1}^{n} 1\big(\max_t\{|y_i(t)|\} > y_{cr}\big)\, 1\big(S_{a,i}(T_0) \in I_k\big), \quad k = 1, \ldots, \bar{k}, \tag{9}$$

where $n_k$ denotes the number of records with response spectra in $I_k$ and $1(A) = 1$ if the $A$ is true and zero otherwise. Note that the fragilities $P_{f,k}(y_{cr})$ are defined in the range $[\min\{S_{a,i}(T_0)\}, \max\{S_{a,i}(T_0)\}]$ and cannot be extended outside this range. The numbers $n_k$ have to be sufficiently large to provide reliable fragility estimates. This requirement is difficult to satisfy in practice due to limited seismic records and in numerical studies due to excessive computational demand.

The triangles and other markers in Fig. 4 are the fragilities $P_{f,k}(y_{cr})$ given by Eq. (9) plotted at the centers of the intervals $\{I_k\}$ of the response spectrum $S_a(T_0)$ with $T_0 = 1$ sec and $\zeta = 5\%$, where $\{y_i(t)\}$ are displacements of the Bouc–Wen oscillator in Eq. (8), $\kappa = 1$, to the ground acceleration records in the reference set. The continuous lines in the figure are lognormal distributions fitted to the probabilities $P_{f,k}(y_{cr})$. The plots show that the information content of even 10,000 ground acceleration records is insufficient to estimate structural performance for large earthquakes as they are rare events. For example, there is just a single SAM cell which yields a non-zero fragility for $y_{cr} = 4.0$.

The current methods for fragility analysis assume that small subsets of the set of available records, referred to as representative ground motions, can characterize site seismicity accurately provided that they are (1) selected in an optimal manner by using objective functions based on responses of linear oscillators and (2) scaled to seismic events of any size. Current selection of representative ground motions is questionable since response maxima of linear oscillators and even simple nonlinear systems are weakly dependent (see Fig. 3). Also, negative effects of scaling are documented extensively [25,26]. The reminder of this study further examines features and limitations of current fragilities.

### 4.2. Reference fragility surfaces

Fragility surfaces are plots of the conditional probabilities $P\big(Y > y_{cr} \mid M, R\big)$ over the $(M, R)$-space. They are constructed for specified soil conditions. As previously, $Y$ denotes a scalar demand parameters, e.g., the maximum response of the Bouc–Wen oscillator in Eq. (8). Note that (1) the conditional probabilities $P\big(Y > y_{cr} \mid M, R\big)$ are uniquely defined for the seismic site and hazard considered in this work and (2) the estimation of the probabilities $P\big(Y > y_{cr} \mid M, R\big)$ requires more samples than that of the fragilities in Section 4.1 to assure that each SAM cell is represented by sufficiently large sets of ground motions. The estimates of the probabilities $P\big(Y > y_{cr} \mid M, R\big)$ are based on $\bar{n} = 1,000$ samples of the ground acceleration processes $X_j(t)$, $j = 1, \ldots, N$, defined by Eqs. (2)–(4).

The following algorithm has been employed to construct reference fragility surfaces.

- *Step 1:* Generate $\bar{n}$ independent samples $\{x_{j,1}(t), \ldots, x_{j,\bar{n}}(t)\}$ of the ground acceleration process $X_j(t)$ and calculate the corresponding responses $\{y_{j,1}(t), \ldots, y_{j,\bar{n}}(t)\}$, $j = 1, \ldots, N$, of, e.g., the Bouc–Wen oscillator in Eq. (8).
- *Step 2:* Estimate $P_{f,j}(y_{cr}) = P\big(Y > y_{cr} \mid (M, R) = (m, r)_j\big)$ from

$$P_{f,j}(y_{cr}) = \frac{1}{\bar{n}}\sum_{i=1}^{\bar{n}} 1\big(\max_t\{|y_{j,i}(t)|\} > y_{cr}\big) \quad j = 1, \ldots, N, \tag{10}$$

  and plot the resulting conditional probabilities $\{P_{f,j}(y_{cr})\}$ at the centers of the SAM cells. The reference fragility surfaces are 2D lognormal distributions fitted to these probabilities.

The panels of Fig. 5 show fragility surfaces for critical levels $y_{cr} = 0.5$, $y_{cr} = 1.0$, $y_{cr} = 2.0$ and $y_{cr} = 4.0$ cm. They are based on $N\,\bar{n} = 300,000$ ground acceleration records, i.e., $\bar{n} = 1000$ records for each of the $N = 300$ SAM cells. The sample size is much larger than that for the fragilities in Fig. 4. Sections through these fragility surfaces are shown in Fig. 6 for $r = 35$ km $m = 6.5$ (left and right panels). The reference fragility surfaces of Fig. 5 show that non-zero fragilities for large critical thresholds, e.g., $y_{cr} = 4.0$ cm, are associated with only a few SAM cells. These cells have small site-to-source distance $r$, large magnitudes $m$, and small probabilities (see the SAM in Fig. 1). Samples associated with these cells are rarely observed even in sets of 10,000 records, which are large relative to commonly available ground acceleration records in Earthquake Engineering.

There is no unique one-to-one mapping between fragility surfaces and fragility curves. Fragility surfaces are failure probabilities conditional on the defining parameters of the ground acceleration process $X(t)$, i.e., the moment magnitude $M$ and the site-to-source distance $R$ for specified soil conditions. Given $(M, R)$ and SAM, the probability law of $X(t)$ is known. In contrast, fragility curves are failure probabilities conditional on IM. The probability law of $X(t)$ cannot be reconstructed from IM. The following algorithm has been used to recast the fragility surfaces of Fig. 5 into the fragility curves of Fig. 7.

- *Step 1:* Calculate the pseudo-spectral response spectra $\{S_{a,j,i}(T)\}$ of the unscaled seismic records $\{x_{j,i}(t)\}$, $i = 1, \ldots, \bar{n}$, $j = 1, \ldots, N$, for a specified reference period $T = T_0$ and damping ratio $\zeta$.
- *Step 2:* Estimate the median of $S_a(T_0)$ and the standard deviation of $\ln\big(S_a(T_0)\big)$ from the set of response spectra $\{S_{a,j,i}(T)\}$. These estimates are denoted by $\mathrm{Med}[S_a(T_0)]$ and $\mathrm{SD}\big[\ln\big(S_a(T_0)\big)\big]$ and are shown in the left and right panels of Fig. 8 for $T_0 = 1$ s and $\zeta = 5\%$. These metrics are standard outputs of the ground-motion prediction equations (GMPEs), which are used subsequently to select ground motions for current fragilities, in lieu of standard GMPEs, for consistency with the hazard model defined in Section 2.
- *Step 3:* Calculate structural responses $\{y_{j,i}(t)\}$, estimate the fragilities in each cell by $P_{f,j}(y_{cr})$ in Eq. (10), and plot them against $\mathrm{Med}[S_a(T_0)]$. These fragilities are shown in Fig. 7 by triangles and other symbols. The continuous lines in Fig. 7 are lognormal distribution fitted to these symbols.

We conclude this subsection with the following three comments. First, it was shown in Fig. 4 that even a set of $n = 10,000$ ground acceleration records, which is large relative to usually available site records, is insufficient to capture extreme rare events, i.e., events from SAM-cells of low probabilities, small $r$'s, and large $m$'s, and that the resulting fragilities are unsatisfactory for large earthquakes. Second, the number of records from individual cells with response spectra larger than specified values $S_a(T_0)$ decreases rapidly with $S_a(T_0)$. For example, this number is 231, 43, and 3 for $S_a(T_0) \geq 150$, 200, and 250 cm/s$^2$ for the SAM-cell $(m, r) = (6.9, 25$ km$)$. Note also that there are only

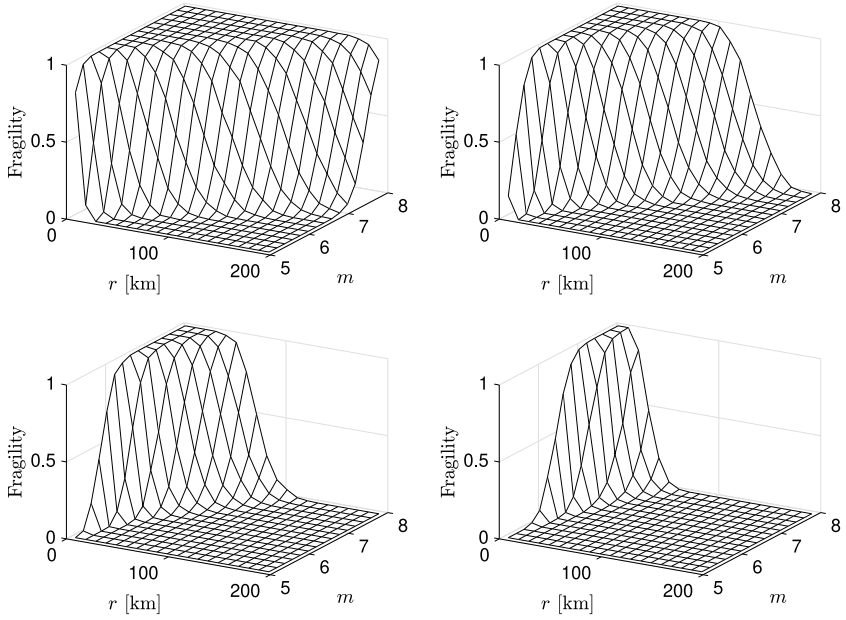


**Fig. 5.** Fragility surfaces for a Bouc–Wen oscillator for $y_{cr} = 0.5$ (top left panel), $y_{cr} = 1.0$ (top right panel), $y_{cr} = 2.0$ (bottom left panel), and $y_{cr} = 4.0$ cm (bottom right panel).

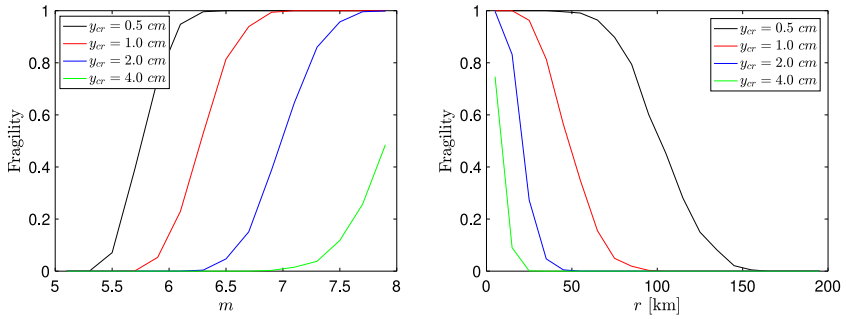


**Fig. 6.** Sections through fragility surfaces for $r = 35$ km (left panel) and $m = 6.5$ (right panel) for the Bouc–Wen system and $y_{cr} = 0.5$, $y_{cr} = 1.0$, $y_{cr} = 2.0$ and $y_{cr} = 4.0$ cm.

11 cells whose records have response spectra larger than $S_a(T_0) \geq 150$ cm/s$^2$, in agreement with the fragility surfaces of Fig. 5. Third, it was possible to construct the fragility surfaces of Fig. 5 and fragility curves of Fig. 7 since the probability law of the seismic ground acceleration is known. These fragilities cannot be constructed in practice because of insufficient ground motion time histories. Also, in contrast to the fragility curves of Fig. 4 which are unsatisfactory for large $y_{cr}$ due to insufficient data, the fragility curves of Fig. 7 are accurate for all critical levels $y_{cr}$. They provide the most accurate representation of structural performance and will be used as reference in the reminder of this study.

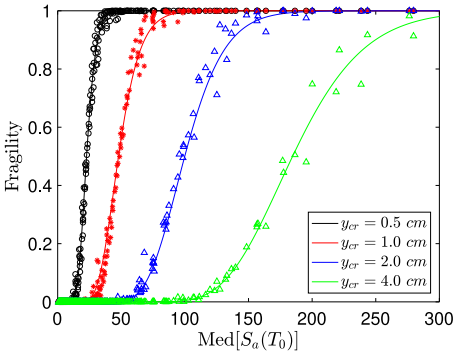


**Fig. 7.** Reference fragility curves derived from the fragility surfaces in Fig. 5 for the Bouc–Wen system for $y_{cr} = 0.5$, $y_{cr} = 1.0$, $y_{cr} = 2.0$ and $y_{cr} = 4.0$ cm.

### 4.3. Current fragility curves

Two types of fragility curves have been developed based on the information provided by the reference set $\{x_1(t), \ldots, x_n(t)\}$ of ground accelerations. The first, referred to as incremental dynamics analysis (IDA) [27,28], selects a single set of ground motions based on an optimization algorithm, scaled them to have the same response spectra at a specified period and damping ratio, and then scaled them to cover desired ranges of spectral responses. The set of ground motion selected for fragility analysis is referred to as the set of ***representative ground motions***. The second, referred to as multiple stripes analysis (MSA) [29,30] partitions the range of desired response spectra into intervals and selects different sets of ground motions in different spectral intervals. The following subsections outline the construction of IDA- and MSA-based fragility curves.

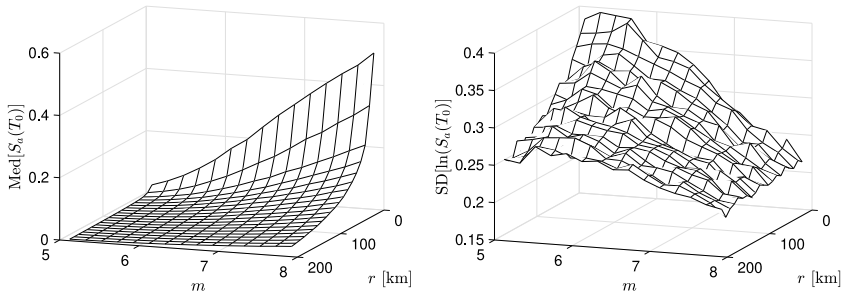


**Fig. 8.** The ground motion prediction model for the SBM, i.e., the median Med$[S_a(T_0)]$ and the standard deviation SD$\left[\ln\left(S_a(T_0)\right)\right]$ for $T_0 = 1$ s and $\zeta = 5\%$ (left and right panels).

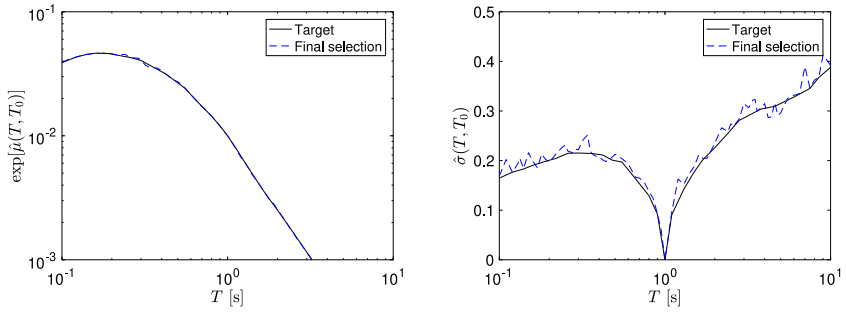


**Fig. 9.** Median $\exp\left(\hat{\mu(T,T_0)}\right)$ and standard deviation $\hat{\sigma}(T,T_0)$ for $T_0 = 1$ s of the conditional spectrum $\ln\left(S_a(T)\right) \mid \ln\left(S_a(T_0)\right)$, see Eq. (12) (solid lines, left and right panels) and estimates from the selected records (dashed lines).

#### 4.3.1. Conditional spectrum

The logarithmic spectrum $\ln\left(S_a(T)\right)$ is assumed to be a non-stationary Gaussian process defined by

$$\ln\left(S_a(T)\right) = \mu_{\ln\left(S_a(T)\right)} + \sigma_{\ln\left(S_a(T)\right)}\, G(T), \quad T \geq 0, \tag{11}$$

where $G(T)$ is a zero-mean, unit-variance Gaussian process with correlation function $\rho(T,T') = E\left[G(T)\,G(T')\right]$. The mean and standard deviation of $\ln\left(S_a(T)\right)$ are denoted by $\mu_{\ln\left(S_a(T)\right)}$ and $\sigma_{\ln\left(S_a(T)\right)}$. For consistency with the hazard model defined in Section 2, all these first- and second-order statistical moments of the process $\ln\left(S_a(T)\right)$ have been calculated using samples of the process $Y(t)$ in Eq. (8), whose input was defined in Eqs. (2) and (3). For reference, the median and standard deviation of $\ln\left(S_a(T)\right)$ have been illustrated in the representation of the empirical GMPE shown in Fig. 8, while the empirical model for the correlation function is similar to the one in [31,32].

The conditional process $\ln\left(S_a(T)\right) \mid \ln\left(S_a(T_0)\right)$ for specified period $T_0$ is referred to as ***conditional spectrum***. Its mean and variance are [22] (Sect. 2.11.5)

$$\begin{aligned}\hat{\mu}(T,T_0) &= \mu_{\ln\left(S_a(T)\right)} + \rho(T,T_0)\left(\ln\left(S_a(T_0)\right)\right. \\ &\quad \left. - \mu_{\ln\left(S_a(T_0)\right)}\right)\sigma_{\ln\left(S_a(T)\right)}/\sigma_{\ln\left(S_a(T_0)\right)} \quad \text{and} \\ \hat{\sigma}(T,T_0)^2 &= \sigma^2_{\ln\left(S_a(T)\right)} - \rho(T,T_0)^2\,\sigma^2_{\ln\left(S_a(T)\right)}\end{aligned} \tag{12}$$

by properties of Gaussian vectors [22]. Note that $\hat{\mu}(T_0,T_0) = \ln\left(S_a(T_0)\right)$ and $\hat{\sigma}(T_0,T_0) = 0$ so that the conditional spectrum coincides with $\ln\left(S_a(T_0)\right)$ since it has no uncertainty at this period. The conditional moments in Eq. (12) with $\mu_{\ln\left(S_a(T)\right)}$ and $\sigma_{\ln\left(S_a(T)\right)}$ in Fig. 8 are used in the subsequent section to select ground motions for fragility analysis.

Under the assumption that $\ln\left(S_a(T)\right)$ are Gaussian variables with means $\mu_{\ln\left(S_a(T)\right)}$ for each $T$, the response spectra $S_a(T) = \exp\left(\ln\left(S_a(T)\right)\right)$ are lognormal variables as exponentials of Gaussian variables. The median of $S_a(T)$ is the exponential of the mean of $\ln\left(S_a(T)\right)$, i.e., $\exp\left(\mu_{\ln\left(S_a(T)\right)}\right)$. The left and right plots of Fig. 8 are estimates of $\exp\left(\mu_{\ln\left(S_a(T)\right)}\right)$ and $\sigma_{\ln\left(S_a(T)\right)}$. Similarly, $\exp\left(\hat{\mu}(T,T_0)\right)$ and $\hat{\sigma}(T,T_0)$ are the median and the standard deviation of the conditional spectrum $\ln\left(S_a(T)\right) \mid \ln\left(S_a(T_0)\right)$.

#### 4.3.2. Ground motion selection

We follow the methodology in [30,33,34] to select representative ground motions from the reference set $\{x_1(t),\ldots,x_n(t)\}$. The methodology has been implemented in a software which is available online [35]. This software was used to find representative ground motions.

Consider first the IDA approach for fragility analysis. In this approach, a single set of records of size $n^* \ll n$ is selected from the reference set $(x_1(t),\ldots,x_n(t))$. The algorithm involves the following two steps.

- *Step 1:* Specify the size $n^*$ of the representative set of ground motions, the period $T_0$, the damping ratio $\zeta$, and the target spectral density $S_a(T)$. The target spectrum is the conditional spectrum of the previous subsection with statistics given by the estimates of Fig. 8.
- *Step 2:* Scale all records in the reference set such that their response spectra at a specified period $T = T_0$ and damping ratio $\zeta$, e.g., $T_0 = 1$ s and $\zeta = 0.05$, will match the ordinate $S_a(T_0)$ of the target spectrum. Denote the resulting scaled records of the reference set by $\{\tilde{x}_1(t),\ldots,\tilde{x}_n(t)\}$. The set of $n^*$ records which are the closest to the target response spectrum in some sense are deemed representative and selected for fragility analysis.

The discrepancy between the response spectra $\{S_{a,1}(T),\ldots,S_{a,n}(T)\}$ of the scaled records $\{\tilde{x}_1(t),\ldots,\tilde{x}_n(t)\}$ and the mean conditional spectrum can be quantified simply by, e.g., the mean square error $\int\left(\mu_{\ln\left(S_a(T)\right)} - S_{a,i}(T)\right)^2 dT$. The quantification of this discrepancies poses notable difficulties if based on the conditional spectrum, since this spectrum is a nonstationary Gaussian process rather than a deterministic function. The software in [35] was used to select representative records in this setting. It was also used to select records for MSA-based fragilities.

The implementation of the IDA approach for selecting ground motions is illustrated in Fig. 9. The solid lines are the median $\exp\left(\hat{\mu}(T,T_0)\right)$ and the standard deviation $\hat{\sigma}(T,T_0)$ provided by Eq. (12). The dashed

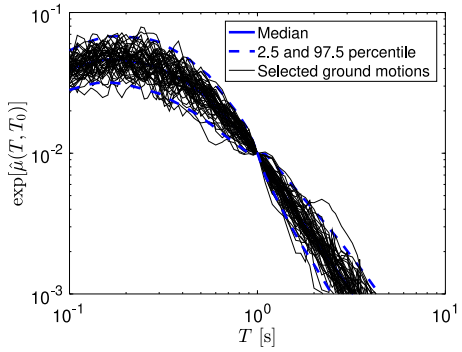


**Fig. 10.** Median of the conditional spectrum $\ln(S_a(T)) \mid \ln(S_a(T_0))$ (heavy solid line), response spectra of selected ground motions (thin lines), and confidence intervals (heavy dashed lines).

lines are the corresponding properties of the selected records. The median $\exp(\hat{\mu}(T,T_0))$ is also shown in Fig. 10 with a heavy solid line. The thin lines are response spectra of the selected ground motion which are scaled to match the target median at $T = T_0$. The heavy dashed lines are confidence intervals.

#### 4.3.3. Fragility curves

Denote by $\{x_{s,1}(t),\dots,x_{s,n^*}(t)\}$, $n^* \ll n$, the set of representative ground motions, i.e., the subset of the reference set $\{x_1(t),\dots,x_n(t)\}$ which has been selected to construct IDA-based fragility curves, and by $\{\tilde{x}_{s,1}(t),\dots,\tilde{x}_{s,n^*}(t)\}$ their scaled versions. The records have been scaled to match the target response spectrum for $T = T_0 = 1$ sec and $\zeta = 5\%$, as illustrated in Figs. 9 and 10.

The construction of IDA-based fragilities involves the following two steps.

- *Step 1:* Calculate structural responses $\{\tilde{y}_{s,r}(t;\xi)\}$ to the ground motions $\{\xi\,\tilde{x}_{s,r}(t)\}$, $r = 1,\dots,n^*$, where the scaling factor $\xi \geq 0$ defines the ground motion intensity, i.e., IM $= \xi$.
- *Step 2:* Estimate fragilities $P_f(\xi)$ by counting the number of responses $\{\tilde{y}_{s,r}(t;\xi)\}$ which exceed critical levels $y_{\rm cr}$, i.e.,

$$\hat{P}_f(\xi) = \frac{1}{n^*}\sum_{r=1}^{n^*} 1\left(\max_t\{|\tilde{y}_{s,r}(t;\xi)|\} > y_{\rm cr}\right), \tag{13}$$

where $1(A) = 1$ if the $A$ is true and zero otherwise. The triangle and other symbols in the left panel of Fig. 11 are the probabilities $\hat{P}_f(\xi)$ plotted at corresponding scaling factors $\xi$. The IDA-fragilities are lognormal distributions fitted to these estimates.

MSA-based fragilities are constructed in a similar manner. First, the range of IMs considered in analysis is partitioned in intervals $\{I_k\}$ and sets of records $\{\tilde{x}_{s,r}^{(k)}(t)\}$, $r = 1,\dots,n^*$, of specified size $n^*$ are selected for these intervals based on conditional spectra corresponding to the midpoints of $\{I_k\}$. The triangle and other symbols in the right panel of Fig. 11 are probabilities calculated as in Eq. (13) based on the ground motions selected for each interval $I_k$ and plotted against the midpoints of these intervals. The MSA-fragilities are lognormal distributions fitted to these estimates.

Consider first the IDA-fragilities in the left panel of Fig. 11 for the Bouc–Wen oscillator of Eq. (8). The solid and dashed lines are IDA-fragilities for target spectra $S_a(T_0) = 0.01$ g and $S_a(T_0) = 0.1$ g and a damping ratio of $\zeta = 5\%$. There are differences between these sets of fragilities and these differences increase with $y_{\rm cr}$. The sensitivity of the IDA-fragilities to the selected value of $S_a(T_0)$ and the fact that there are no criteria for selecting $S_a(T_0)$ place doubts on the reliability of IDA-based estimates of structural performance.

Consider now the MSA-fragilities in the right panel of Fig. 11 for the same Bouc–Wen system. The figure shows four families of fragilities corresponding to various values of $y_{\rm cr}$. The solid, dashed, and dotted lines in each families are for $n^* = 10$, 50, and 100 selected samples. The MSA-based fragilities are also sensitive to the particular parameters used in analysis. For example, they seem to stabilize as $n^*$ increases for $y_{\rm cr} = 4$ but diverge as $n^*$ increases for $y_{\rm cr} = 2$. As a result, different analyst will obtain different fragilities depending on the particular parameters selected for analysis. This is a matter of great concern for practice.

## 5. Current and reference fragilities

The characterization of the seismic performance of structural systems poses significant difficulties because the physics of earthquakes is partially understood and the sets of recorded ground accelerations are small. The plots of Fig. 4 show that even sets of records which are large relative to typically available data in Earthquake Engineering are insufficient to construct fragilities for moderate and large seismic events. The construction of these fragilities requires information beyond the available sets of seismic ground acceleration records.

There are at least two directions to overcome these difficulties and construct fragilities. The first is to (1) view seismic ground acceleration records as samples of stochastic processes whose statics can be inferred from physics and data [9] and (2) use samples of these processes to develop, e.g., fragility surfaces [5], which do not use surrogates, e.g., response spectra or other IMs, to characterize site seismicity. This approach was used to construct the reference fragility curves of Figs. 4 and 7 and the reference fragility surface of Fig. 5.

The second approach, represented by the IDA and MSA methods, increases the information content of available set of ground acceleration records by scaling subsets of the set of ground motions which is selected by optimization criteria. Simplicity and capability of delivering fragilities for any range of IMs are the main feature of the IDA and MSA fragilities. The heuristic arguments employed to implement these fragilities are matters of concerns. This section examines potential negative effects of some of these arguments on the current fragility curves.

***Intensity measures:*** We have seen that the spectral acceleration is an unsatisfactory metric for the seismic hazard. It was shown in Fig. 3 that this IM provides limited if any information on the performance of even simple nonlinear systems. Moreover, the predictive capability of response spectra decreases with earthquake magnitude. This means that fragilities defined as probabilities that demand parameters exceed critical values conditional IMs are unreliable estimates of structural performance particularly for large seismic events.

***Scaling ground motions:*** Generally, the number of seismic ground acceleration records decreases with the earthquake magnitude and so does the accuracy of the resulting fragility estimates. Scaling can be viewed as a heuristic approach to augment the information content of the set of seismic ground motions selected for fragility analysis. It delivers the same number of records at all IMs of interest.

Yet, scaling has unfavorable side effects. For example, scaling (1) generates records which have the same frequency content for small and large earthquakes which is at variance with physics and observations [9,15], (2) produces biased estimates of structural responses [26], and (3) changes the law of the ground acceleration process and these changes may result in conservative or unconservative estimates of structural performance [25]. We note that, in contrast to scaling, the frequency content of the ground accelerations generated by the seismological model in [9] depends on earthquake magnitude, see the spectral densities in Fig. 2.

***Dynamics:*** The representative records consists of the members of the site seismic ground acceleration records which are the closest in some

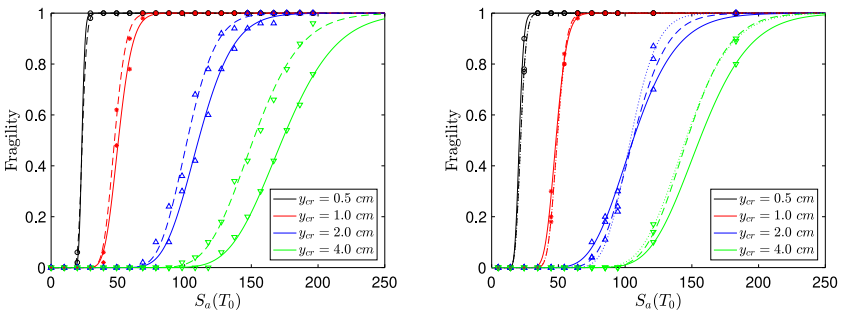


**Fig. 11.** IDA-fragilities (left panel) for $S_a(T_0) = 0.01$ g and $S_a(T_0) = 0.1$ g (solid and dashed lines) and MSA-fragilities (right panel) for $n^* = 10$, 50, and 100 (solid, dashed, and dotted lines).

sense to target response spectra $S_a(T)$. The selection of the representative records does not account for the mechanical properties of structural systems which largely determine their dynamic responses. As a result, the same sets of representative records are selected irrespective of the dynamical properties of the structural system under consideration. This is a severe limitation since the set of representative seismic records must account for system properties. For example, the scatter plots of Fig. 3 show that ground motions which cause large displacements in linear oscillators may or may not yield large displacements of the (nonlinear) Bouc–Wen system. This means that ground motions which are representative for a system may not be representative for another system.

***Statistics:*** The estimation of the probability of rare events require large data sets. For example, suppose that $p(z) = P(Z > z) \sim O(10^{-5})$ is the quantity of interest, where $Z$ is a random variable. Accurate Monte Carlo (MC) estimates of $p(z)$ require at least $10^6$ samples of $Z$. *Importance sampling* (IS) is an alternative approach which allows to estimate $p(z)$ from relative small sets of samples. In this approach, the estimates of $p(z)$ are constructed from samples of $Z$ under a biasing distribution of $Z$ rather than its nominal distribution [36]. The resulting estimates of $p(z)$ are unbiased. However, depending on the biasing distribution, the variances of IS estimates can be smaller or larger than those of MC estimates for the same sample size.

The fragility estimates delivered by the IDA and MSA methods are based on small sets of samples selected from available populations of ground acceleration records, i.e., the reference set of $n = 10{,}000$ records in our study. The algorithms for record selection can be viewed as a heuristic importance sampling procedure which does not attempt to minimize the variance of the estimated quantity of interest, i.e., structural fragilities. This observation suggests that alternative selection procedure, e.g., random selection of representative records, may yield similar fragilities.

The above intuition is supported by the plots of Fig. 12. The dot-dashed and dashed lines, and the thin gray solid lines are the reference fragility curve of Fig. 7, the MSA-fragility, $n^* = 100$, of Fig. 11, and IDA-fragilities based on random selection for $y_{cr} = 2.0$ cm. The family of IDA-fragilities has 50 members and each member is constructed from 50 ground motions selected at random from the reference set of records. Once a set of 50 records is selected, the IDA-methodology is applied to construct fragilities. We note that the IDA-fragilities based on randomly selected records band the MSA-fragilities and that some members of this family of fragilities are closer to the reference fragility curve than the MSA-fragility.

***Reference and current fragilities:*** The plots of Fig. 13 include reference and current fragilities for $y_{cr} = 0.5$ cm and $y_{cr} = 1.0$ cm (left and right top panels) and $y_{cr} = 2.0$ cm and $y_{cr} = 4.0$ cm (left and right bottom panels). The dot-dashed and solid lines are the reference fragilities of Figs. 4 and 7. The dotted and dashed lines are the IDA-fragilities of Fig. 11 (left panel) and the MSA-fragilities Fig. 11 (right panel). For low critical thresholds $y_{cr}$, there are relatively small differences between

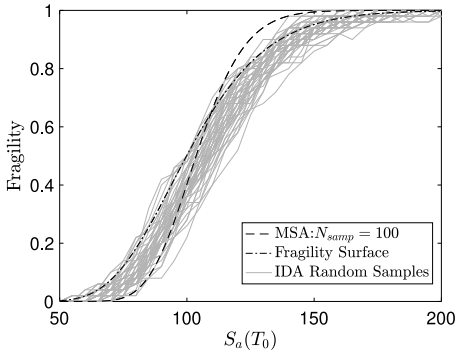


**Fig. 12.** Reference fragility of Fig. 7, MSA-fragility of Fig. 11, and IDA-fragilities based on random sampling (heavy dot-dashed, heavy dashed, and thin solid lines).

various fragilities (see top panels). These differences increase with $y_{cr}$ since the corresponding fragilities are controlled by large seismic events (see bottom panels). As previously mentioned, the reference fragilities of Fig. 4 are unsatisfactory for large $y_{cr}$ due to limited data. The IDA-fragilities are at variance with the reference fragilities of Fig. 7 and the MSA-fragilities. The differences between these fragilities increase with the critical threshold $y_{cr}$. MSA-fragilities constitute a notable improvement over IDA-fragilities. Yet, they differ from the reference fragilities based on fragility surfaces and these differences increase with $y_{cr}$. They under- and over-estimate the reference fragilities depending on the intensity of seismic events.

## 6. Conclusions

A hypothetical seismic site has been constructed for which the law of the seismic ground acceleration process $X(t)$ is specified. This setting allows to assess the performance of IDA- and MSA-fragility curves without ambiguity since truth is known. The site seismicity was represented by a reference set of 10,000 independent samples of $X(t)$. This set is larger than most available sets of site seismic ground acceleration records. It was used to implement the IDA- and MSA-fragilities and assess their performance.

It was found that (1) even 10,000 ground acceleration records are insufficient to estimate structural performance for large seismic events if records are not scaled, (2) fragility surfaces, i.e., failure probabilities conditional on the defining parameters of the seismic ground acceleration process $X(t)$, are the only rigorous estimates of structural performance but they require large data sets, (3) IDA- and MSA-fragilities are sensitive to the particular parameters used for their construction, (4) ground motion selection algorithms are of questionable value since fragilities based on samples delivered by these algorithms and randomly selected ground motions are similar, (5) MSA-fragilities are superior to IDA-fragilities, (6) differences between MSA- and reference fragilities increase with ground motion intensity, and

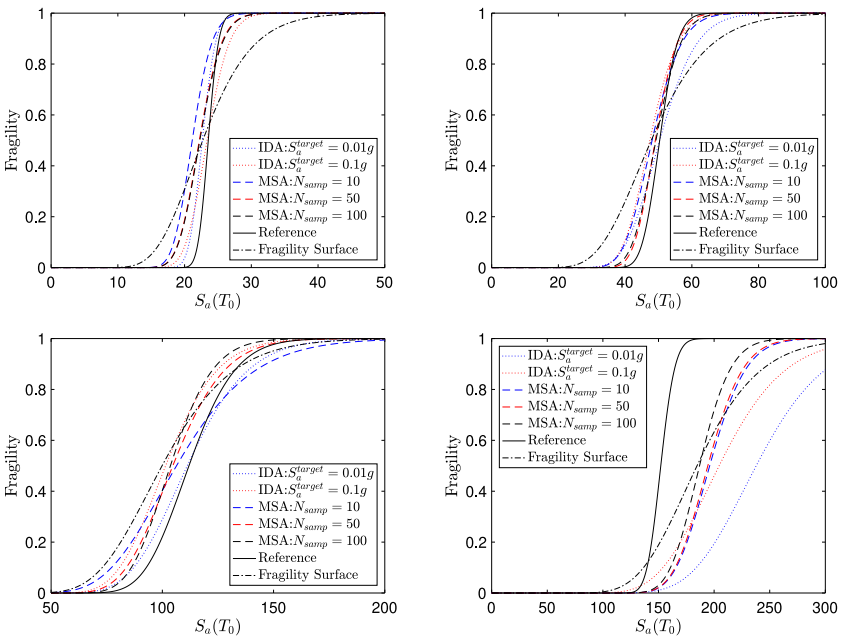


**Fig. 13.** Reference and current fragilities for the Bouc–Wen system for $y_{cr} = 0.5$ cm and $y_{cr} = 1.0$ cm (left and right top panels) and $y_{cr} = 2.0$ cm and $y_{cr} = 4.0$ cm (left and right bottom panels).

(7) MSA-fragilities seem to converge/diverge with the sample size depending on values of critical demand parameters.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: M. Grigoriu & A. Radu

## Acknowledgment

The work reported in this paper has been partially supported by the National Science Foundation, USA under grants CMMI-1265511 and CMMI-1639669. This support is gratefully acknowledged.

## References

[1] H. Ebrahimian, F. Jalayer, A. Lucchini, F. Mollaioli, G. Manfredi, Preliminary ranking of alternative scalar and vector intensity measures of ground shaking, Bull. Earthq. Eng. (2015) http://dx.doi.org/10.1007/s10518-015-9755-9.

[2] P. Tothong, N. Luco, Probabilistic seismic demand analysis using advanced ground motion intensity measures, Earthq. Eng. Struct. Dyn. 36 (2007) 1837–1860, http://dx.doi.org/10.1002/eqe.696.

[3] M. Ciano, M. Gioffré, M. Grigoriu, The role of intensity measures on the accuracy of seismic fragilities, Probab. Eng. Mech. 60 (2020) 103041.

[4] M. Grigoriu, Do seismic intensity measures (IMs) measure up? Probab. Eng. Mech. 46 (2016) 80–93.

[5] C. Kafali, M. Grigoriu, Seismic fragility analysis: Application to simple linear and nonlinear systems, Earthq. Eng. Struct. Dyn. 36 (2007) 1885–1900.

[6] W. Baker, J. Conditional mean spectrum: Tool for ground-motion selection, J. Struct. Eng. 137 (2011) 322–331.

[7] N.S. Kwong, A.K. Chopra, R.K. McGuire, Evaluation of ground motion selection and modification procedures using synthetic ground motions, Earthq. Eng. Struct. Dyn. 44 (2015) 1841–1861, http://dx.doi.org/10.1002/eqe.2558.

[8] B.A. Bradley, L.S. Burks, J.W. Baker, Ground motion selection forsimulation-based seismic hazard and structural reliability assessment, Earthq. Eng. Struct. Dyn. 44 (13) (2015) 2321–2340, http://dx.doi.org/10.1002/eqe.2588.

[9] A. Papageorgiou, Engineering seismology, in: D.E. Beskos, E. Kausel (Eds.), Advances in Earthquake Engineering, 1997.

[10] D. Sornette, A. Sornette, General theory of the modified Gutenberg–Richter law for large seismic moments, Bull. Seismol. Soc. Am. 89 (4) (1999) 1121–1130.

[11] USGS, U.S. Geological Survey: 2009 Earthquake probability mapping, 2017, Last checked on 02 November 2017. URL https://geohazards.usgs.gov/eqprob/2009/.

[12] USGS, U.S. Geological Survey: Uniform hazard tool, 2018, Last checked on 11 May 2018. URL https://earthquake.usgs.gov/hazards/interactive/.

[13] A. Papageorgiou, K. Aki, A specific barrier model for the quantitative description of inhomoeneous faulting and the prediction of strong ground motion. applications of the model, Bull. Seismological Soc. Am. 73 (4) (1983) 953–978.

[14] A. Papageorgiou, K. Aki, A specific barrier model for the quantitative description of inhomoeneous faulting and the prediction of strong ground motion. description of the model, Bull. Seismological Soc. Am. 73 (3) (1983) 693–722.

[15] B. Halldorsson, A.S. Papageorgiou, Calibration of the specific barrier model to earthquakes of different tectonic regions, Bull. Seismilogical Soc. Am. 95 (4) (2005) 1276–1300.

[16] A. Radu, M. Grigoriu, A site-specific ground-motion simulation model: Application for Vrancea earthquakes, Soil Dyn. Earthq. Eng. 111 (2018) 77–86.

[17] A. Radu, M. Grigoriu, A site-specific seismological model for probabilistic seismic-hazard assessment, Bull. Seismol. Soc. Am. 104 (6) (2014) 3054–3071, http://dx.doi.org/10.1785/0120140013.

[18] S. Rezaeian, A.D. Kiureghian, Simulation of orthogonal horizontal ground motion components for specified earthquake and site characteristics, Earthq. Eng. Struct. Dyn. 41 (2) (2012) 335–353.

[19] A. Tsioulou, A. Taflanidis, C. Galasso, Modification of stochastic ground motion models for matching target intensity measures, Earthq. Eng. Struct. Dyn. (2017) 1–22.

[20] C. Vlachos, K.G. Papakonstantinou, G. Deodatis, Predictive model for site specific simulation of ground motions based on earthquake scenarios, Earthq. Eng. Struct. Dyn. 47 (1) (2017) 195–218.

[21] K. Goda, C. Petrone, R. De Risi, T. Rossetto, Stochastic coupled simulation of strong motion and tsunami for the 2011 Tohoku, Japan earthquake, Stoch. Environ. Res. Risk Assess. (2016) 1–19.

[22] M. Grigoriu, Applications in Science and Engineering, Birkhäuser, Boston, 2002.

[23] N.L. Johnson, S. Kotz, Discrete Distributions, John Wiley & Sons, Inc., New York, 1969.

[24] A. Radu, M. Grigoriu, An earthquake-source-based metric for seismic fragility analysis, Bull. Earthq. Eng. (2018).

[25] M. Grigoriu, To scale or not to scale seismic ground records, J. Eng. Mech. 137 (4) (2011) 284–293.

[26] N. Luco, P. Bazzurro, Does amplitude scaling of ground motion records result in biased nonlinear structural drift responses? Earthq. Eng. Struct. Dyn. 36 (2007) 1813–1835, http://dx.doi.org/10.1002/eqe.695.

[27] D. Vamvatsikos, C. Cornell, Incremental dynamic analysis Earthquake, Eng. Struct. Dyn. 31 (2002) 491–514.

[28] C. Goulet, C. Haselton, J. Mitrani-Reiser, J. Beck, G. Deierlein, K. Porter, J. Stewart, Assessment of probability of collapse and design for collapse safety, Earthq. Eng. Struct. Dyn. 36 (2007) 1973–1997.

[29] J. Baker, Efficient analytical fragility function fitting using dynamic structural analysis, Earthq. Spectra 31 (1) (2015) 579–599.

[30] T. Lin, C. Haselton, J. Baker, Conditional spectrum-based ground motion selection. part I: Hazard consistency for risk-based assessments, Earthq. Eng. Struct. Dyn. 42 (12) (2013) 1847–1865.

[31] T. Lin, C. Haselton, J. Baker, Conditional spectrum-based ground motion selection. part I: Hazard consistency for risk-based assessments, Earthq. Eng. Struct. Dyn. 42 (12) (2013) 1847–1865.
[32] J.W. Baker, N. Jayaram, Correlation of spectral acceleration values from NGA ground motion models, Earthq. Spectra 24 (1) (2008) 299–317.
[33] N. Jayaram, T. Lin, J.W. Baker, A computationally efficient ground-motion selection algorithm for matching a target response spectrum mean and variance, Earthq. Spectra 27 (3) (2011) 797–815.
[34] J.W. Baker, C. Lee, An improved algorithm for selecting ground motions to match a conditional spectrum, J. Earthq. Eng. 22 (4) (2018) 708–723.
[35] J.W. Baker, C. Lee, Ground-motion selection software, 2019, https://github.com/bakerjw/CS_Selection, last (Accessed:24 November 2019).
[36] R. Srinivasan, Importance sampling, in: Applications in Communications and Detection, Springer, New York, 2002.