Earthquake Engineering & Structural Dynamics
The Journal of the International Association for Earthquake Engineering

EARTHQUAKE ENGINEERING & STRUCTURAL DYNAMICS
Earthquake Engng Struct. Dyn. 2017; **46**:1711–1723
Published online 21 February 2017 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/eqe.2883

SHORT COMMUNICATION

Assessing uncertainty in estimation of seismic response for PBEE

Iunio Iervolino*,†

Dipartimento di Strutture per l'Ingegneria e l'Architettura, Università degli Studi di Napoli Federico II, Naples, Italy

SUMMARY

State-of-the-art approaches to probabilistic assessment of seismic structural reliability are based on simulation of structural behavior via nonlinear dynamic analysis of computer models. Simulations are carried out considering samples of ground motions supposedly drawn from specific populations of signals virtually recorded at the site of interest. This serves to produce samples of structural response to evaluate the failure rate, which in turn allows to compute the failure risk (probability) in a time interval of interest. This procedure alone implies that uncertainty of estimation affects the probabilistic results. The latter is seldom quantified in risk analyses, although it may be relevant. This short paper discusses some basic issues and some simple statistical tools, which can aid the analyst towards the assessment of the impact of sample variability on fragility functions and the resulting seismic structural risk. On the statistical inference side, the addressed strategies are based on consolidated results such as the well-known delta method and on some resampling plans belonging to the bootstrap family. On the structural side, they rely on assumptions and methods typical in performance-based earthquake engineering applications. Copyright © 2017 John Wiley & Sons, Ltd.

Received 25 October 2016; Revised 16 January 2017; Accepted 22 January 2017

KEY WORDS: fragility; statistical inference; bootstrap; delta method; performance-based earthquake engineering

1. INTRODUCTION

Performance-based earthquake engineering or PBEE [1] assumes that the occurrence of the (main) earthquakes at the site of the construction can be described by a homogeneous Poisson process (HPP). It can be shown that, in the framework of PBEE, if seismic damage accumulation is not considered (e.g., [2]), also the process of earthquakes causing structural failure is a HPP. This is convenient, because the HPP is a one-parameter model, that is, the failure probability in any time interval only depends on the failure rate, λ_f. It is usually computed via Equation (1), where $P[f|im]$ is a function providing the failure probability conditional to the values of a ground motion intensity measure (IM), that is, the *fragility* of the structure, and $d\lambda_{im} = d(im) \cdot d\lambda_{im}/d(im)$ is obtained from the derivative of the seismic *hazard curve*. The latter, usually provided by engineering seismologists, associates to each IM value, *im*, the rate of earthquakes exceeding it at the site where the structure is located, λ_{im}.

$$\lambda_f = \int_{im} P[f|im] \cdot |d\lambda_{im}| \tag{1}$$

In the state-of-the-art approach, the fragility can be evaluated via several procedures that are all based on nonlinear dynamic analysis of a structural numerical model. The simulations aim at the

*Correspondence to: Iunio Iervolino, Dipartimento di Strutture per l'Ingegneria e l'Architettura, Università degli Studi di Napoli Federico II, via Claudio 21, 80125, Naples, Italy.
†E-mail: iunio.iervolino@unina.it

generation of samples of structural response given samples of ground motions, selected consistently with the seismic hazard of the site [3]. It immediately follows that only an estimate of the failure rate of the structure is obtained. To better understand this issue, one can consider that a probability model is often assigned to define structural fragility, lognormal being a typical choice. Consequently, the failure rate is provided by Equation (2), where $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF) and $\{\eta, \beta\}$ are parameters.

$$\lambda_f(\eta, \beta) = \int_{im} \Phi\left[\frac{\log(im) - \eta}{\beta}\right] \cdot |d\lambda_{im}| \tag{2}$$

In this context, given the hazard curve, the failure rate is a function of $\{\eta, \beta\}$. These parameters are typically not known, and are estimated based on a sample of ground motions. Because of *record-to-record variability* of structural response, their estimated values, say $\left\{\hat{\eta}, \hat{\beta}\right\}$, are expected to change when changing the sample used, the ground motion selection procedure, and all other elements of the analysis staying the same. Therefore, also $\hat{\lambda}_f = \lambda_f\left(\hat{\eta}, \hat{\beta}\right)$ varies with the sample. In other words, $\hat{\lambda}_f$ is an estimator (i.e., function of the sample) of the fixed, yet unknown, failure rate, λ_f. As such, $\hat{\lambda}_f$ can be regarded as a random variable (RV), the distribution of which quantifies the uncertainty involved in the risk estimation. To evaluate the mean and the standard error [4] of the failure rate estimator, $\sigma_{\hat{\lambda}_f}$, is the goal of this simple study.

It is to note that, in fact, uncertainty of estimation in fragility assessment may also arise from structural modeling. Moreover, several components needed to evaluate the hazard curve appearing in Equation (1) may suffer from uncertainty of estimation. More in general, all the factors contributing to the seismic risk of structures in PBEE, that is *hazard*, *fragility*, and *loss*, are based on several models the calibration of which relies on estimations; see for example [5]. To limit the scope of the study, the focus herein is on record-to-record variability of structural response for fragility assessment and the effect on the resulting failure rates.

The remainder of the paper is structured such that the case of lognormal fragility is addressed first. Then, some non-parametric and parametric resampling plans, belonging to the *bootstrap* family [6], are illustrated. Finally, uncertainty of estimation when the failure rate is computed via the *Cornell method* [7] is addressed. Examples equip the study to illustrate how the discussed strategies can be employed to get information about the estimator of seismic structural reliability.

2. COMMON FRAGILITY MODELING STRATEGIES

To briefly recall the most common strategies to fit fragility functions, it is assumed that a suite of records (n in number) is available. It may be a single set of records, which is manipulated (i.e., scaled) to get structural responses spanning the IM domain; for example, via incremental dynamic analysis or IDA [8]. It may also be the case of multiple sets of records, each of which is used at a specific IM level; that is, multi-stripe analysis (MSA) that is considered generally superior to IDA in warranting consistency with hazard. The following assumes, for simplicity, that response data are from IDA, also for those situations in which MSA could be employed [9].

2.1. IM-based approach

A simple and effective way to evaluate the structural fragility is the so-called IM-based approach, in which, typically, IDA is employed to obtain a sample of ground motion IMs that caused failure of the structural numerical model, taking for each record in the set the minimal IM-value at which an undesired structural response is observed. See for example Figure 1 (left) where the IM is the spectral acceleration, $Sa(T)$, at an oscillation period (T) equal to 1.82 s and 5% damping. The engineering demand parameter (EDP), measuring structural response, is the maximum inter-story drift ratio (IDR), and the failure threshold is $edp_f = 0.03$ (see Section 3.3 for structural details).

DOI: 10.1002/eqe

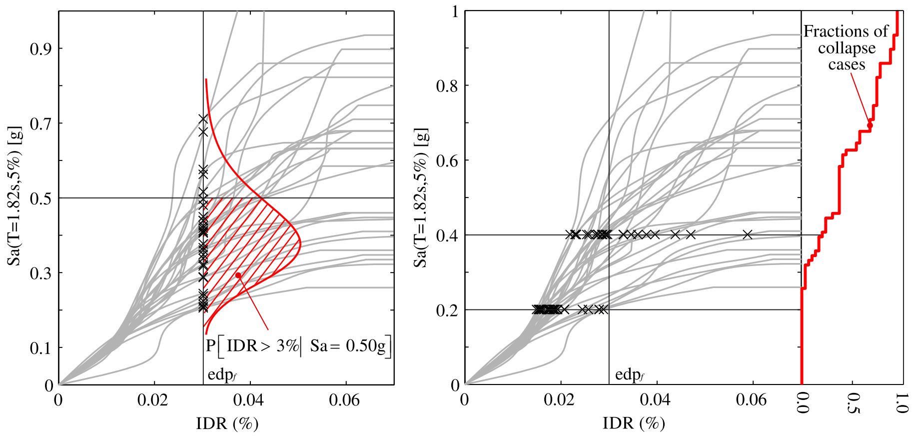

Figure 1. Example of IM-based and EDP-based fragility assessment, left and right respectively, from incremental dynamic analysis using 30 records (see Section 3.3 for details). Crosses represent individual structural responses. IDR is inter-story drift ratio. [Colour figure can be viewed at wileyonlinelibrary.com]

Such a sample of n values, $\underline{im} = \{im_1, im_2, \ldots, im_n\}$, is considered drawn from the distribution of the ground motion intensity causing failure of the structure, say IM_f. The CDF of this RV is the seismic fragility. As discussed with respect to Equation (2), it is a common assumption in PBEE applications to assign the lognormal model to it, Equation (3). The parameters represent the mean and the standard deviation of the logarithms of the IM (RV) causing structural failure. Therefore, based on the $\underline{im}$ sample, they can be estimated via Equations (4).

$$P[f|im] = P\left[IM_f \leq im\right] \approx \Phi\left(\hat{\eta}, \hat{\beta}\right) \tag{3}$$

$$\begin{cases} \hat{\eta} = \frac{1}{n} \cdot \sum_{i=1}^{n} \log(im_i) \\ \hat{\beta}^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} \left[\log(im_i) - \hat{\eta}\right]^2 \end{cases} \tag{4}$$

2.2. *Truncated IDA*

It is also to mention the case in which it is not possible to evaluate the IM causing the failure of the structure for all n records in the ground motion set. For example, when IDA is intentionally carried out until a $im_{\max}$ level and not beyond. This may be because larger IMs may require scaling records by large factors, which is believed to sometimes cause bias in the evaluation of structural response, or simply because analyzing the structure at large IM levels may be computationally demanding, yet not particularly impacting the risk assessment, which is dominated by less large yet more frequently exceeded IMs, as Equation (1) suggests.

In these situations, it may be that failure is observed only for a fraction of the records, say m in number; for the other $n-m$, the information is only that failure occurs for intensity larger than $im_{\max}$. This means that the sample of IM_f is *censored* and the parameters of the lognormal fragility can be estimated by the maximum likelihood criterion in Equation (5),* where $\phi(\cdot)$ is the standard normal probability density function (PDF).

*The equation is from [9], yet revised.

$$\left\{\hat{\eta},\hat{\beta}\right\} = \underset{\eta,\beta}{\arg\max}\left[\sum_{i=1}^{m}\left(\log\left\{\frac{1}{im_i\cdot\beta}\cdot\phi\left[\frac{\log(im_i)-\eta}{\beta}\right]\right\}\right)+(n-m)\cdot\log\left\{1-\Phi\left[\frac{\log(im_{\max})-\eta}{\beta}\right]\right\}\right] \tag{5}$$

2.3. EDP-based approach

Alternative to the IM-based is the EDP-based approach, which works at fixed IM levels. At $IM = im$, all the records in the ground motion set feature the same intensity, and simulations provide a vector, $\underline{edp} = \{edp_1, edp_2, \ldots, edp_n\}$, representing structural response conditional to im. The $\underline{edp}$ vector can be partitioned in two: the EDP-values that determine failure, say k, and those for which failure has not occurred, $n-k$.

Applying this procedure to a number of intensity levels im_i, $i = \{1, 2, \ldots, m\}$, yields data to fit a fragility model [9]. Under this hypothesis, $\left\{\hat{\eta},\hat{\beta}\right\}$ can be obtained via the binomial likelihood of Equation (6), where k_i and $(n-k_i)$ are the number of failures and non-failures observed at $IM = im_i$, respectively. Figure 1 (right) provides an example, where two IM levels are highlighted, 0.2 g and 0.4 g. Crosses beyond edp_f represent failure.

$$\left\{\hat{\eta},\hat{\beta}\right\} = \underset{\eta,\beta}{\arg\max}\left[\sum_{i=1}^{m}\left(\log\binom{n}{k_i}+k_i\cdot\log\left\{\Phi\left[\frac{\log(im_i)-\eta}{\beta}\right]\right\}+(n-k_i)\cdot\log\left\{1-\Phi\left[\frac{\log(im_i)-\eta}{\beta}\right]\right\}\right)\right] \tag{6}$$

It is also to mention that at a specific intensity level, and for a specific record, it may be that the structure is failed, but the EDP-value is not available from the computer simulation; these situations are referred to as *collapse cases* [10]. To represent this kind of failure for all IM levels, the fractions of collapse cases over the total number of records ($n = 30$) are given as the thick curve in Figure 1. The total number of failures at each IM level, k_i, is given by the sum of EDPs beyond edp_f and the collapse cases.

2.4. Non-parametric fragility

It is functional to the following discussions to recall that the fragility of the structure is in principle unknown. In other words, referring for example to the IM-based approach, the $\{im_1, im_2, \ldots, im_n\}$ values in the $\underline{im}$ vector are realizations of independent- and identically-distributed RVs, the distribution being the fragility of the structure, the exact shape of which is not available. As discussed, this issue is often addressed by assuming a probabilistic model whose parameters are calibrated based on $\underline{im}$. In principle, an alternative, which does not require a hypothesis about the underlying model, is to build the empirical CDF of the intensity causing structural failure. The empirical fragility is a stepwise function defined between minus and plus infinity as in Equation (7), where $I_{(im_i \le im)}$ is an indicator function that equals 1 if $im_i \le im$ and 0 otherwise.

$$P[f|im] = \frac{1}{n}\cdot\sum_{i=1}^{n} I_{(im_i \le im)} \tag{7}$$

3. UNCERTAINTY OF ESTIMATION IN LOGNORMAL FRAGILITY

3.1. Known distribution of parameter estimators

If the estimates of the parameters of the lognormal fragility function are obtained via Equations (4), their estimators have known distributions. In particular, the sampling mean is normally distributed,

DOI: 10.1002/eqe

while the rescaled estimator of the variance is chi-squared-distributed with $n-1$ degrees of freedom; Equations (8).[†] Because these distributions are known, and because the RVs they refer to are also statistically-independent [4], then the mean and the variance of the failure rate estimator can be obtained solving the integrals of Equations (9), where $f(\cdot)$ indicates PDF.

$$\begin{cases} \hat{\eta} \sim N\left(\eta, \dfrac{\beta^2}{n}\right) \\ \dfrac{n-1}{\beta^2} \cdot \hat{\beta}^2 \sim \chi^2_{n-1} \end{cases} \tag{8}$$

$$\begin{cases} E\left[\hat{\lambda}_f\right] = \int_0^{+\infty} \int_{-\infty}^{+\infty} \hat{\lambda}_f(u,v) \cdot f_{\hat{\eta}}(u) \cdot f_{\hat{\beta}^2}(v) \cdot du \cdot dv \\ VAR\left[\hat{\lambda}_f\right] = \int_0^{+\infty} \int_{-\infty}^{+\infty} \left\{\hat{\lambda}_f(u,v) - E\left[\hat{\lambda}_f\right]\right\}^2 \cdot f_{\hat{\eta}}(v) \cdot f_{\hat{\beta}^2}(u) \cdot du \cdot dv \end{cases} \tag{9}$$

Note that the distributions in Equations (8) depend on the true parameters of the fragility, which are not available and estimates are used instead. Therefore, Equations (9) also provide estimates.

3.2. Delta method

An approximation of the mean and the variance of the estimator of the failure rate can also be obtained by the so-called *delta method* (e.g., [11]), which is based on the Taylor series expansion of the function $\hat{\lambda}_f = \lambda_f\left(\hat{\eta}, \hat{\beta}\right)$, Equations (10). To illustrate how it works, it is applied to the lognormal fragility, even if it was discussed that this case can be solved as described in the previous section when the parameters are estimated via Equations (4).

$$\begin{cases} E\left[\hat{\lambda}_f\right] \approx \hat{\lambda}_f + \dfrac{1}{2} \cdot VAR[\hat{\eta}] \cdot \dfrac{\partial^2 \hat{\lambda}_f}{\partial \hat{\eta}^2} + \dfrac{1}{2} \cdot VAR\left[\hat{\beta}^2\right] \cdot \dfrac{\partial^2 \hat{\lambda}_f}{\partial \left(\hat{\beta}^2\right)^2} \\ VAR\left[\hat{\lambda}_f\right] \approx VAR[\hat{\eta}] \cdot \left(\dfrac{\partial \hat{\lambda}_f}{\partial \hat{\eta}}\right)^2 + VAR\left[\hat{\beta}^2\right] \cdot \left[\dfrac{\partial \hat{\lambda}_f}{\partial \left(\hat{\beta}^2\right)}\right]^2 \end{cases} \tag{10}$$

Rewriting the failure rate in Equation (2) as in Equation (11), showing the lognormal PDF, helps to obtain the partial derivatives given in Equations (12). Because the variances of the estimators are also available, see Equations (13)[‡] descending from Equations (8), the variance and the mean of the estimator can be obtained applying Equations (10). Note that all these equations need to be derived only once, and can always be applied. They allow to evaluate the mean and the variance of $\hat{\lambda}_f$ by only computing two integrals each.

$$\lambda_f = \int_0^{+\infty} \int_0^{im} \frac{1}{u \cdot \sqrt{2 \cdot \pi \cdot \beta^2}} \cdot e^{-\frac{1}{2} \cdot \left[\frac{\log(u) - \eta}{\beta}\right]^2} \cdot du \cdot |d\lambda_{im}| \tag{11}$$

[†]For simplicity, herein both point estimates of the parameters and their estimators (i.e., the random variables) will be indicated with the same symbol.

[‡]The about equal symbol means that the variance estimate is used in lieu of the true (unknown) value.

DOI: 10.1002/eqe

$$
\begin{cases}
\dfrac{\partial \hat{\lambda}_f}{\partial \hat{\eta}} = \int\limits_0^{+\infty} \int\limits_0^{im} \dfrac{\log(u) - \hat{\eta}}{\hat{\beta}^2} \cdot \dfrac{e^{-\frac{1}{2}\left[\frac{\log(u) - \hat{\eta}}{\hat{\beta}}\right]^2}}{u \cdot \sqrt{2 \cdot \pi \cdot \hat{\beta}^2}} \cdot du \cdot |d\lambda_{im}| \\[2ex]
\dfrac{\partial^2 \hat{\lambda}_f}{\partial \hat{\eta}^2} = \int\limits_0^{+\infty} \int\limits_0^{im} \left\{ \left[\dfrac{\log(u) - \hat{\eta}}{\hat{\beta}^2} \right]^2 - \dfrac{1}{\hat{\beta}^2} \right\} \cdot \dfrac{e^{-\frac{1}{2}\left[\frac{\log(u) - \hat{\eta}}{\hat{\beta}}\right]^2}}{u \cdot \sqrt{2 \cdot \pi \cdot \hat{\beta}^2}} \cdot du \cdot |d\lambda_{im}| \\[2ex]
\dfrac{\partial \hat{\lambda}_f}{\partial \left(\hat{\beta}^2\right)} = \int\limits_0^{+\infty} \int\limits_0^{im} \left\{ \dfrac{[\log(u) - \hat{\eta}]^2}{2 \cdot \hat{\beta}^4} - \dfrac{1}{2 \cdot \hat{\beta}^2} \right\} \cdot \dfrac{e^{-\frac{1}{2}\left[\frac{\log(u) - \hat{\eta}}{\hat{\beta}}\right]^2}}{u \cdot \sqrt{2 \cdot \pi \cdot \hat{\beta}^2}} \cdot du \cdot |d\lambda_{im}| \\[2ex]
\dfrac{\partial^2 \hat{\lambda}_f}{\partial \left(\hat{\beta}^2\right)^2} = \int\limits_0^{+\infty} \int\limits_0^{im} \left\{ \dfrac{[\log(u) - \hat{\eta}]^4}{4 \cdot \hat{\beta}^8} - \dfrac{3 \cdot [\log(u) - \hat{\eta}]^2}{2 \cdot \hat{\beta}^6} + \dfrac{3}{4 \cdot \hat{\beta}^4} \right\} \cdot \dfrac{e^{-\frac{1}{2}\left[\frac{\log(u) - \hat{\eta}}{\hat{\beta}}\right]^2}}{u \cdot \sqrt{2 \cdot \pi \cdot \hat{\beta}^2}} \cdot du \cdot |d\lambda_{im}|
\end{cases}
\tag{12}
$$

$$
\begin{cases}
VAR[\hat{\eta}] \approx \dfrac{\hat{\beta}^2}{n} \\[2ex]
VAR\left[\hat{\beta}^2\right] \approx \dfrac{2 \cdot \hat{\beta}^4}{n - 1}
\end{cases}
\tag{13}
$$

3.3. *Example*

To illustrate the aforementioned procedures, the fragility analysis of a case study structure is considered. It is a four-story, three-bay, steel moment resisting perimeter frame. It belongs to the archetype structures designed and used for the purposes of the NIST GCR 10-917-8 report [12]. The numerical model is a 2D centerline idealization implemented in the OpenSees analysis platform [13]. Material nonlinearity is accounted for by means of a lumped-plasticity approach, where the properties of the plastic hinges at the element edges are estimated using the regression equations suggested by [14]. Geometric nonlinearities ($P - \Delta$ effects) are also considered.

The structure is supposedly located in the town of Ancona (lon. 13.515; lat. 43.614), in central Italy (Figure 2, left), on a site characterized by a shear-wave velocity in the upper 30 m equal to 270 m/s. The hazard for the site has been computed using the seismic source model also used in [15] and the ground motion prediction equation of [16]. The hazard curve

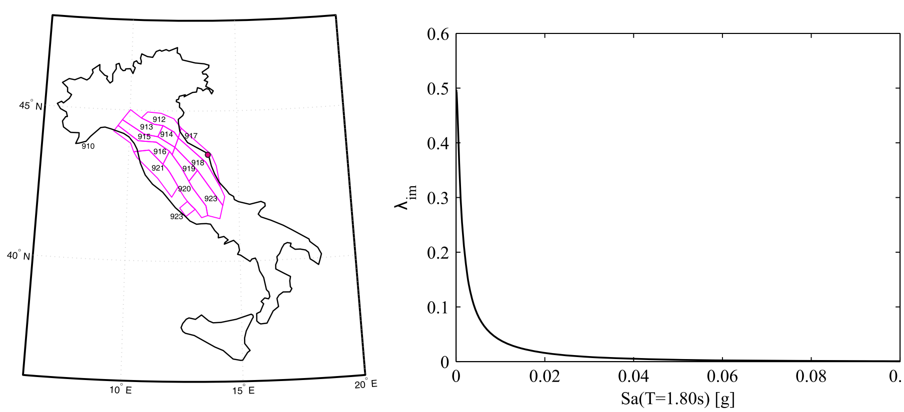

Figure 2. Ancona site (dot) and relevant seismic sources (left); hazard curve (right). [Colour figure can be viewed at wileyonlinelibrary.com]

DOI: 10.1002/eqe

Table I. Mean and standard error of the failure rate in the case of lognormal distribution and approximated distributions/variances of its parameter estimators.

$\hat{\lambda}_f$ [events/year]	Method	$E\left[\hat{\lambda}_f\right]$ [events/year]	$\sigma_{\hat{\lambda}_f} = \sqrt{VAR\left[\hat{\lambda}_f\right]}$ [events/year]	Coefficient of variation [−]
$2.50 \cdot 10^{-5}$	*Approx. distrib.*	$2.56 \cdot 10^{-5}$	$6.34 \cdot 10^{-6}$	25%
	Delta	$2.56 \cdot 10^{-5}$	$6.28 \cdot 10^{-6}$	25%

is in terms of annual rate of exceedance of spectral pseudo-acceleration at T = 1.80s (Figure 2, right).[§]

The seismic input for the analyses is a set of 30 single-component recorded accelerograms selected among the records of the FEMA P695 [17] far-field set and the *Engineering Strong Motion database* (http://esm.mi.ingv.it/). All records are from free-field, firm soil conditions, with the nearest distance of the recording stations to the fault plane being in the 7–33 km range and correspond to causal events of magnitude ranging from 6.0 to 7.5. Only one horizontal component per station is used. Figure 1 shows the IDA results for the model subjected to the records.

In this application, the IM-based approach of Section 2.1 is considered to get the failure rate of the structure and then uncertainty of estimation. Data indicating failure are the 30 IM-values where individual IDA curves cross IDR equal to $edp_f = 0.03$ (cross-marked data of Figure 1, left). The estimators of Equations (4) are used to get the fragility curve, they result equal to $\left\{\hat{\eta} = -0.99, \hat{\beta} = 0.34\right\}$; that is, the mean and standard deviation of the logarithms of the IM causing failure. Fragility is integrated with the hazard via Equation (1), yielding an annual failure rate equal to $2.50 \cdot 10^{-5}$ [events/year] for the structure at the site of interest.[**]

At this point, to get the mean and the variance of $\hat{\lambda}_f$, integrals in Equations (9) are solved numerically. For comparison, also the delta method of Equations (10) is applied. The results for both are given in Table I. In this example, the approximated delta method yields very similar results with respect to the solution via Equations (9), and the coefficient of variation of the failure rate is about 25%.

4. RESAMPLING SEISMIC FRAGILITY

Beyond the simple case described in the previous section, it may be that the distributions of the fragility parameter estimators are not readily available. In these situations, to get a sense of the uncertainty of estimation, some techniques collectively known as *bootstrap*, and typically requiring Monte Carlo simulation, can be an option. The basic version of bootstrap relies on the generation of samples from an empirical distribution obtained from data regarding a phenomenon of interest; alternatively, a parametric model calibrated on data can be used for resampling. Owing to their apparent simplicity, these techniques are routinely used in research, including earthquake engineering (e.g., [18, 19]).

In this section, the basics of bootstrap and parametric bootstrap are recalled so that they can be used to approximate $\sigma_{\hat{\lambda}_f}$. However, bootstrap can be applied to more general problems; for example, it can be used to evaluate the distribution of the estimator. On the other hand, it is also the subject of a large deal of statistics literature and, therefore, the interested reader is referred to dedicated readings about this and other resampling plans; for example, [6].

[§]The 1.82 s spectral acceleration hazard is herein considered approximately equal to the 1.80 s one.

[**]This rate is about equal to the annual failure probability, P_f, because for such a small rate $P_f(1yr) = 1 - e^{-\lambda_f} \approx \lambda_f$.

DOI: 10.1002/eqe

4.1. Resampling

Given the discussion in the previous section, the failure rate in Equation (1) can be seen as a function of the response data used to fit the fragility function. In other words, referring for example to the IM-based approach: $\hat{\lambda}_f = \lambda_f(im_1, im_2, \ldots, im_n)$.

The bootstrap-based standard deviation, $\hat{\sigma}_{\hat{\lambda}_f}$, is obtained via Monte Carlo simulation, sampling the empirical fragility function discussed in Section 2.4. In practical terms, sampling from the empirical distribution means resampling data from $\underline{im} = \{im_1, im_2, \ldots, im_n\}$, that is, random drawing with replacement to obtain several other vectors still of size n. A generic vector obtained by resampling $\underline{im}$ can be indicated as $\underline{im}^* = \{im_1^*, im_2^*, \ldots, im_n^*\}$. The key point is that each $\underline{im}^*$ can be considered as a new realization of the fragility function. Consequently, it can be employed to estimate the failure rate of the structure. If this procedure is repeated n_b number of times, $\hat{\sigma}_{\hat{\lambda}_f}$ can be computed via Equation (14), where $\hat{\lambda}_{f,i}^*$ is the failure rate from the i-th resampling and $\hat{\mu}_{\hat{\lambda}_f^*} = \frac{1}{n_b} \cdot \sum_{i=1}^{n_b} \hat{\lambda}_{f,i}^*$.

$$\hat{\sigma}_{\hat{\lambda}_f} = \sqrt{\frac{1}{n_b - 1} \cdot \sum_{i=1}^{n_b} \left(\hat{\lambda}_{f,i}^* - \hat{\mu}_{\hat{\lambda}_f^*} \right)^2} \tag{14}$$

4.2. Parametric resampling

Beyond the non-parametric case, there may be situations in which the fragility function has a probabilistic model associated to it, yet the distributions of the parameter estimators are not readily available. The *parametric bootstrap* may be a solution to this issue.

To illustrate it, let us continue with the IM-based example above in which the $\underline{im} = \{im_1, im_2, \ldots, im_n\}$ data vector is used to fit a lognormal model via Equation (5). The parametric bootstrap, consists of sampling an arbitrary number of vectors of the kind $\underline{im}^* = \{im_1^*, im_2^*, \ldots, im_n^*\}$ from the lognormal distribution, which features the parameters estimated from the original sample, that is, $\Phi\left(\hat{\eta}, \hat{\beta}\right)$. If it is decided to truncate IDA when the first m records out of n induce structural failure, then the $\underline{im}^*$ vector is ordered and the first m intensity values are taken to estimate $\left\{\hat{\eta}^*, \hat{\beta}^*\right\}$ via Equation (5), such that a new fragility function, $\Phi\left(\hat{\eta}^*, \hat{\beta}^*\right)$, is used to integrate with hazard and obtain $\hat{\lambda}_f^*$ (this means that in each run $im_{\max}$ is the m-th ordered IM). Repeating this procedure multiple times allows to compute Equation (14) to get the parametric-bootstrap-based $\hat{\sigma}_{\hat{\lambda}_f}$.

An analogous procedure can be applied for the EDP-based approach, as it is illustrated via the examples in the following section.

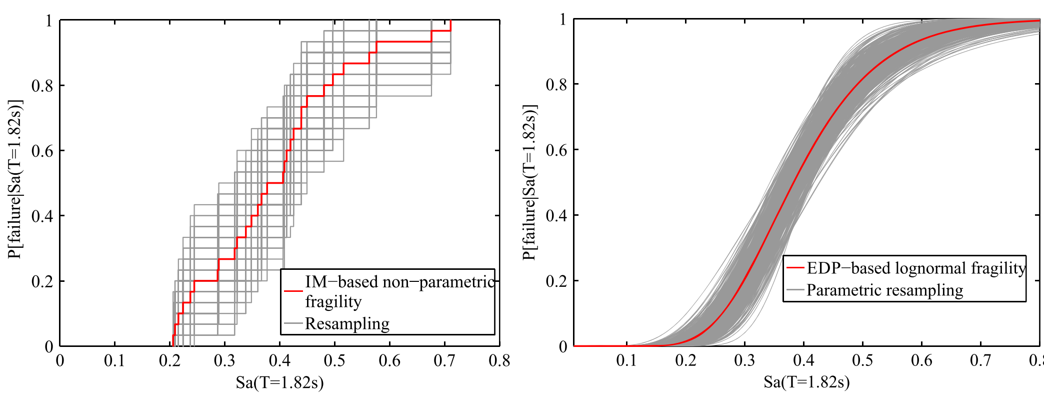

Figure 3. Fragilities from 500 resampling of the IM-based empirical fragility (left); fragilities from 500 samplings of the lognormal EDP-based fragility data (right). [Colour figure can be viewed at wileyonlinelibrary.com]

DOI: 10.1002/eqe

Table II. Mean and standard error the failure rate with 500 non-parametric or parametric resamplings.

$\hat{\lambda}_f$ [events/year]	Approach	$\hat{\mu}_{\hat{\lambda}_f^*}$ [events/year]	$\hat{\sigma}_{\hat{\lambda}_f}$ [events/year]	Coefficient of variation [−]
$2.58 \cdot 10^{-5}$	*IM-based non-parametric*	$2.62 \cdot 10^{-5}$	$5.17 \cdot 10^{-6}$	20%
$1.98 \cdot 10^{-5}$	*EDP-based parametric*	$2.00 \cdot 10^{-5}$	$3.99 \cdot 10^{-6}$	20%
$2.52 \cdot 10^{-5}$	*IM-based (truncated) parametric*	$2.56 \cdot 10^{-5}$	$6.48 \cdot 10^{-6}$	25%

4.3. Examples

To apply bootstrap of Section 4.1, the vector of 30 IM-values is used to build the empirical fragility curve (thick line in Figure 3, left). Five-hundred samples are extracted from it, such that a new vector of 30 IM-values is obtained each time, and a new empirical fragility function is built in every run. The 500 individual fragility curves are identified by thin lines in Figure 3 (left). Each curve is, then, numerically integrated with the hazard as per Equation (1) yielding 500 failure rates. The failure rate computed with the original empirical CDF is given in Table II, along with the non-parametric-bootstrap-based mean and standard error of the failure rate estimator.

To apply parametric resampling of Section 4.2, the EDP-based approach of Section 2.3 is applied. Five IM levels are considered, $Sa(T=1.82s)=\{0.2g, 0.3g, 0.4g, 0.5g, 0.6g\}$. Applying Equation (6) to the resulting data, when the failure threshold is $edp_f=0.03$, yields a lognormal fragility with parameters $\left\{\hat{\eta} = -0.96, \hat{\beta} = 0.30\right\}$; the obtained fragility is the thick curve in Figure 3 (right). At each of the five IM levels, a sample of 30 failure/no-failure cases is drawn from a binomial distribution with $p_i = \Phi\left\{\left[\log(im_i) - \hat{\eta}\right]/\hat{\beta}\right\}$, $i=\{1, 2, \ldots, 5\}$, and the parameters of the fragility function are re-calibrated via Equation (6). The obtained fragilities in 500 runs are given in Figure 3 (right) as thin lines, while the results in terms of failure rates are given in Table II.

To illustrate the parametric resampling procedure in the case of truncated IDA of Section 2.2, the lognormal fragility curve is calibrated via Equation (5). This fragility, shown in Figure 4 (left) as the thick line, was obtained censoring IDA at the IM level where the first 15 out of 30 records induce failure, which yields $\left\{\hat{\eta} = -0.98, \hat{\beta} = 0.36\right\}$. When resampling, 30 IM-values are extracted in each run. The first 15 values of the extracted IMs are kept (i.e., $im_{\max}$ is the median of the extracted IMs). These values, in turn, are used to re-fit a lognormal fragility via Equation (5). The fragilities from the simulations are given in Figure 4 (left) as thin lines. In Figure 4 (right), the distribution of $im_{\max}$ is also shown. The mean and the standard error of the failure rate are given in Table II.

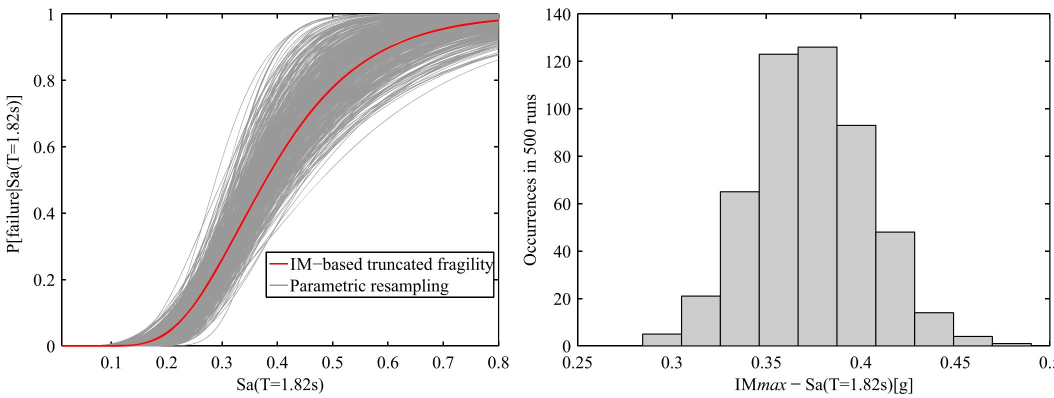

Figure 4. Truncated IM-based lognormal fragilities (left) and distribution of truncation IM to observe 15 out of 30 failures in each IDA (right). [Colour figure can be viewed at wileyonlinelibrary.com]

DOI: 10.1002/eqe

5. UNCERTAINTY OF ESTIMATION IN THE CORNELL METHOD

The Cornell method [7, 20] to assess seismic structural reliability is based on the *cloud* analysis that is a less computationally demanding alternative, with respect to IDA and MSA, to evaluate the relationship between ground motion and structural response. In particular, a suite of n unscaled ground motion records, with arbitrary intensities, $\{im_1, im_2, \dots, im_n\}$, is used to obtain a sample of structural responses, $\{edp_1, edp_2, \dots, edp_n\}$. A log-linear regression model, of the type in Equation (15), is then calibrated based on these data. In the equation, $\left\{\hat{a}, \hat{b}\right\}$ are the estimated coefficients and ε is the residual, that is, in the classical case, a zero mean Gaussian RV, with estimated standard deviation $\hat{\beta}_D = \sqrt{(n-2)^{-1} \cdot \sum_{i=1}^{n} \left[\log(edp_i) - \left(\hat{a} + \hat{b} \cdot \log(im_i)\right)\right]^2}$.

$$\log(edp) = \hat{a} + \hat{b} \cdot \log(im) + \varepsilon \tag{15}$$

This model is used to obtain the failure rate as per Equation (16), where edp_f is the median capacity, and β_C is the corresponding logarithmic standard deviation under lognormality assumption. The parameters $\{k_0, k\}$ are obtained from the linearization of the hazard curve in the log–log space around the rate corresponding to the IM-value obtained transforming edp_f via the linear regression. The cloud regression enters in the equation via: $\hat{\alpha} = e^{\hat{a}}$, $\hat{b}$, and $\hat{\beta}_D$.

$$\hat{\lambda}_f \approx k_0 \cdot \left(\frac{edp_f}{\hat{\alpha}}\right)^{-\frac{k}{\hat{b}}} \cdot e^{\frac{1}{2} \frac{k^2}{\hat{b}^2} \cdot \left(\hat{\beta}_D^2 + \beta_C^2\right)} \tag{16}$$

The analytical format of the failure rate allows to apply the delta-method to propagate the uncertainty in the evaluation of the structural response, and to obtain the mean and standard deviation of the failure rate estimator. In fact, Cornell's equation can be interpreted as a function transforming the estimators of the IM-EDP relationship in the estimator of the failure rate: $\hat{\lambda}_f = \lambda_f\left(\hat{a}, \hat{b}, \hat{\beta}_D\right)$. From the Taylor series expansion, it follows that the mean and the variance of $\hat{\lambda}_f$ can be approximated via Equations (17), requiring the partial derivatives of the estimator with respect to the RVs involved and their variances/covariances, similar to Equations (10). (Note that, under classical hypotheses, only $COV\left[\hat{a}, \hat{b}\right]$ is not equal to zero among the covariances of the involved estimators.)

$$\begin{cases} E\left[\hat{\lambda}_f\right] \approx \hat{\lambda}_f + \frac{1}{2} \cdot VAR[\hat{a}] \cdot \frac{\partial^2 \hat{\lambda}_f}{\partial \hat{a}^2} + \frac{1}{2} \cdot VAR\left[\hat{b}\right] \cdot \frac{\partial^2 \hat{\lambda}_f}{\partial \hat{b}^2} + \frac{1}{2} \cdot VAR\left[\hat{\beta}_D^2\right] \cdot \frac{\partial^2 \hat{\lambda}_f}{\partial \left(\hat{\beta}_D^2\right)^2} + COV\left[\hat{a}, \hat{b}\right] \cdot \frac{\partial^2 \hat{\lambda}_f}{\partial \hat{a} \cdot \partial \hat{b}} \\ VAR\left[\hat{\lambda}_f\right] \approx VAR[\hat{a}] \cdot \left(\frac{\partial \hat{\lambda}_f}{\partial \hat{a}}\right)^2 + VAR\left[\hat{b}\right] \cdot \left(\frac{\partial \hat{\lambda}_f}{\partial \hat{b}}\right)^2 + VAR\left[\hat{\beta}_D^2\right] \cdot \left[\frac{\partial \hat{\lambda}_f}{\partial \left(\hat{\beta}_D^2\right)}\right]^2 + 2 \cdot COV\left[\hat{a}, \hat{b}\right] \cdot \frac{\partial \hat{\lambda}_f}{\partial \hat{a}} \cdot \frac{\partial \hat{\lambda}_f}{\partial \hat{b}} \end{cases} \tag{17}$$

It can be recognized that the first-order and second-order partial derivatives needed are those in Equations (18). In the case of linear regression, the moments of the estimators are also known [4]. They are recalled in Equations (19) for convenience. In these equations, $\hat{\eta} = \frac{1}{n} \cdot \sum_{i=1}^{n} \log(im_i)$ is the mean of the IMs of the records used for cloud analysis. As it happened for the delta method applied to the lognormal fragility in Section 3.2, all these equations need to be derived only once, and can always be applied.

DOI: 10.1002/eqe

$$\begin{cases} \dfrac{\partial\hat{\lambda}_f}{\partial\hat{a}} = \dfrac{k}{\hat{b}}\cdot\hat{\lambda}_f \\ \dfrac{\partial^2\hat{\lambda}_f}{\partial\hat{a}^2} = \dfrac{k}{\hat{b}}\cdot\dfrac{\partial\hat{\lambda}_f}{\partial\hat{a}} \\ \dfrac{\partial\hat{\lambda}_f}{\partial\hat{b}} = \left[\dfrac{k}{\hat{b}^2}\cdot(\log C - \hat{a}) - \dfrac{k^2}{\hat{b}^3}\cdot\left(\hat{\beta}_D^2 + \beta_C^2\right)\right]\cdot\hat{\lambda}_f \\ \dfrac{\partial^2\hat{\lambda}_f}{\partial\hat{b}^2} = \left[-\dfrac{2\cdot k}{\hat{b}^3}\cdot(\log C - \hat{a}) + \dfrac{3\cdot k^2}{\hat{b}^4}\cdot\left(\hat{\beta}_D^2 + \beta_C^2\right)\right]\cdot\hat{\lambda}_f + \left[\dfrac{k}{\hat{b}^2}\cdot(\log C - \hat{a}) - \dfrac{k^2}{\hat{b}^3}\cdot\left(\hat{\beta}_D^2 + \beta_C^2\right)\right]\cdot\dfrac{\partial\hat{\lambda}_f}{\partial\hat{b}} \\ \dfrac{\partial^2\hat{\lambda}_f}{\partial\hat{a}\cdot\partial\hat{b}} = \dfrac{k}{\hat{b}^2}\cdot(\log C - \hat{a})\cdot\dfrac{\partial\hat{\lambda}_f}{\partial\hat{a}} - \dfrac{k}{\hat{b}^2}\cdot\hat{\lambda}_f - \dfrac{k^2}{\hat{b}^3}\cdot\left(\hat{\beta}_D^2 + \beta_C^2\right)\cdot\dfrac{\partial\hat{\lambda}_f}{\partial\hat{a}} \\ \dfrac{\partial\hat{\lambda}_f}{\partial\left(\hat{\beta}_D^2\right)} = \dfrac{1}{2}\cdot\dfrac{k^2}{\hat{b}^2}\cdot\hat{\lambda}_f \\ \dfrac{\partial^2\hat{\lambda}_f}{\partial\left(\hat{\beta}_D^2\right)^2} = \dfrac{1}{2}\cdot\dfrac{k^2}{\hat{b}^2}\cdot\dfrac{\partial\hat{\lambda}_f}{\partial\left(\hat{\beta}_D^2\right)} \end{cases} \tag{18}$$

$$\begin{cases} VAR[\hat{a}] \approx \left\{\dfrac{1}{n} + \dfrac{\hat{\eta}^2}{\sum_{i=1}^{n}[\log(im_i) - \hat{\eta}]^2}\right\}\cdot\hat{\beta}_D^2 \\ VAR\left[\hat{b}\right] \approx \dfrac{\hat{\beta}_D^2}{\sum_{i=1}^{n}[\log(im_i) - \hat{\eta}]^2} \\ COV\left[\hat{a},\hat{b}\right] \approx \dfrac{-\hat{\eta}\cdot\hat{\beta}_D^2}{\sum_{i=1}^{n}[\log(im_i) - \hat{\eta}]^2} \\ VAR\left[\hat{\beta}_D^2\right] \approx \dfrac{2\cdot\hat{\beta}_D^4}{n-2} \end{cases} \tag{19}$$

5.1. Example

As an example, the Cornell method is applied to the structure introduced in Section 3.3 also using the same records. In particular, the structural responses from the IDAs when the scaling factor of each of the 30 records is equal to one 1 were considered to carry out the linear regression in Equation (15); see Figure 5 (left). The estimates are: $\left\{\hat{a} = \text{-}2.55,\ \hat{b} = 0.88,\ \hat{\beta}_D = 0.25\right\}$, $\left\{VAR[\hat{a}] = 0.0195,\ VAR\left[\hat{b}\right] = 0.0027,\ COV\left[\hat{a},\hat{b}\right] = 0.0069,\ VAR\left[\hat{\beta}_D^2\right] = 0.00028\right\}$. The other parameters required to apply Equation (16) are $k_0 = 4.65\cdot10^{-7}$ and $k = 3.41$ from the linearization of the hazard curve (Figure 5, right). Then, because the capacity was considered deterministic (i.e., $\beta_C = 0$) for simplicity, no other information is needed to compute $\hat{\lambda}_f = 3.03\cdot10^{-5}$. At this point, applying Equations (17) to (19) one obtains $E\left[\hat{\lambda}_f\right] = 3.15\cdot10^{-5}$ and $VAR\left[\hat{\lambda}_f\right] = 9.43\cdot10^{-11}$, yielding a coefficient of variation equal to 31%.[††]

[††]Beyond the cloud analysis, the difference of the results of the Cornell method with respect to other cases also depends on the log-linear approximation of the hazard curve. For discussions, the interested reader can refer, for example, to [21] and references therein.

DOI: 10.1002/eqe

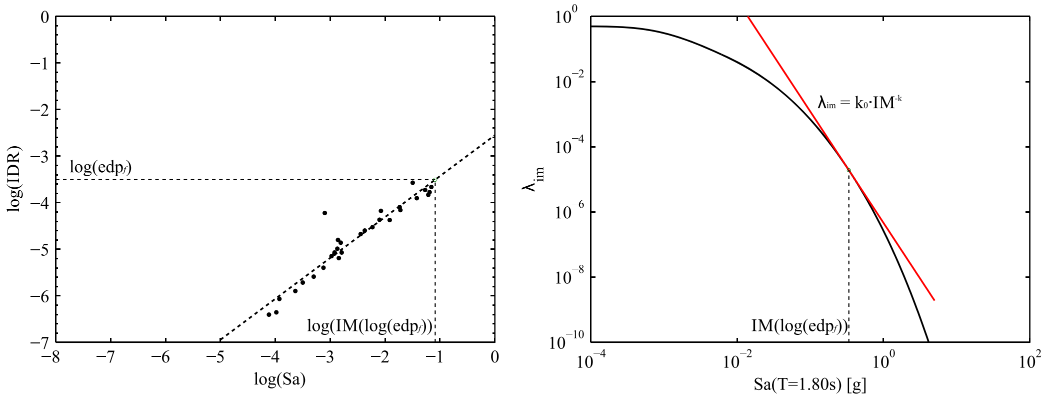

Figure 5. Cloud analysis and linear regression (left); hazard curve and linear approximation around the acceleration corresponding to the structural capacity (right). [Colour figure can be viewed at wileyonlinelibrary.com]

6. FINAL REMARKS

At least because the procedures to assess seismic structural risk employ simulation of dynamic response with ground motion samples, the failure rate (probability) is affected by uncertainty of estimation. In this short study, a few easy-to-apply techniques have been considered to get a quantitative sense of such an uncertainty, when it descends from record-to-record-variability of response.

Three categories of procedures have been described that are especially suited in the case of incremental dynamic or multi-stripe analyses: one which is applicable when the (approximate) distribution of the estimator of the fragility parameters is available, one based on the delta method approximation, and one based on bootstrap. In this context, quantification of uncertainty in estimation was also obtained for the Cornell method, which consists of an approximate closed-form solution for the failure rates based on a log-linear hazard curve and cloud-type structural analysis.

The procedures described can be of aid to the structural analysts to provide the risk estimate with uncertainty of estimation, which in turn allows more informative comparisons of reliability between structures and/or better evaluation of probabilistic seismic losses.

ACKNOWLEDGMENTS

The study was developed in the framework the ReLUIS-DPC 2014-2018 project, funded by the Italian *Civil Protection Department*. The discussions with professor Massimiliano Giorgio (*Seconda Università degli Studi di Napoli*, Italy) were very important. Dr. Georgios Baltzopoulos (*Istituto di Tecnologie della Costruzione*, CNR, Italy) contributed to the examples based on a structural model provided by Dr. Dimitrios Vamvatsikos (*National Technical University of Athens*, Greece). The reviews from Prof. Jack W. Baker (*Stanford University*, USA) and Prof. Paolo Franchin (*Sapienza Università di Roma*, Italy) are also gratefully acknowledged.

REFERENCES

1. Cornell CA, Krawinkler H. Progress and challenges in seismic performance assessment. *PEER Center News* 2000, **3**(2):1–3.
2. Iervolino I, Giorgio M, Chioccarelli E. Markovian modeling of seismic damage accumulation. *Earthquake Engineering and Structural Dynamics* 2016, **45**(3):441–461.
3. Cornell CA. Hazard analysis, ground motions, and probabilistic assessment for PBSD. Proc. of *International Workshop on Performance Based Design*, Bled, Slovenia, 2004.
4. Mood AM, Graybill FA, Boes DC. *Introduction to the Theory of Statistics*. McGraw-Hill: New York, US, 1974.
5. Bradley BA. A critical examination of seismic response uncertainty analysis in earthquake engineering. *Earthquake Engineering and Structural Dynamics* 2013, **42**(11):1717–1729.
6. Efron B. The jackknife, the bootstrap, and other resampling plans. *CBMS-NSF, Regional Conference Series in Applied Mathematics*. Society for industrial and applied mathematics: Philadelphia, US, 1982.
7. Cornell CA. Calculating building seismic performance reliability: a basis for multi-level design norms. Proc. of *11th World Conference on Earthquake Engineering, 11WCEE*, 1996.

DOI: 10.1002/eqe

8. Vamvatsikos D, Cornell CA. Applied incremental dynamic analysis. *Earthquake Spectra* 2004, **20**(2):523–553.
9. Baker JW. Efficient analytical fragility function fitting using dynamic structural analysis. *Earthquake Spectra* 2015, **31**(1):579–599.
10. Shome N, Cornell CA. Structural seismic demand analysis: Consideration of "Collapse". In *8th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability*. University of Notre Dame, US, 2000.
11. Oehlert GW. A note on the delta method. *The American Statistician* 1992, **46**(1):27–29.
12. NIST. *Evaluation of the FEMA P-695 methodology for quantification of building seismic performance factors,* Report NIST GCR 10-917-8. Prepared for the US National Institute of Standards and Technology by the NEHRP Consultants Joint Venture, Gaithersburg, MD, US, 2010.
13. Mazzoni S, McKenna F, Scott MH, Fenves GL. *OpenSees Command Language Manual*. Pacific Earthquake Engineering Research (PEER) Center: CA, US, 2006 [http://opensees.berkeley.edu/].
14. Lignos DG, Krawinkler H. Deterioration modeling of steel components in support of collapse prediction of steel moment frames under earthquake loading. *Journal of Structural Engineering, ASCE* 2011, **137**(11):1291–1302.
15. Iervolino I, Chioccarelli E, Convertito V. Engineering design earthquakes from multimodal hazard disaggregation. *Soil Dynamics and Earthquake Engineering* 2011, **31**(9):1212–1231.
16. Akkar S, Bommer JJ. Empirical equations for the prediction of PGA, PGV, and spectral accelerations in Europe, the Mediterranean region, and the Middle East. *Seismological Research Letters* 2010, **81**(2):195–206.
17. FEMA P695. *Quantification of Building Seismic Performance Factors*. Rep. FEMA P695, Federal Emergency Management Agency: Washington, DC, US, 2009.
18. Baker JW, Cornell CA. Vector-valued intensity measures incorporating spectral shape for prediction of structural response. *Journal of Earthquake Engineering* 2008, **12**(4):534–554.
19. Bradley BA. Epistemic uncertainties in component fragility functions. *Earthquake Spectra* 2010, **26**(1):41–62.
20. Cornell CA, Jalayer F, Hamburger RO, Foutch DA. Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines. *Journal of Structural Engineering, ASCE* 2002, **128**(4):526–533.
21. Vamvatsikos D. Derivation of new SAC/FEMA performance evaluation solutions with second-order hazard approximation. *Earthquake Engineering and Structural Dynamics* 2013, **42**(8):1171–1188.

DOI: 10.1002/eqe