Received: 1 October 2018 | Accepted: 17 November 2018

DOI: 10.1002/eqe.3145

RESEARCH ARTICLE

WILEY

On the number of records for structural risk estimation in PBEE

Georgios Baltzopoulos[1] | Roberto Baraschino[1,2] | Iunio Iervolino[1]

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

[2]Department of Civil Engineering, National Technical University of Athens, Athens, Greece

Correspondence
Georgios Baltzopoulos, Università degli Studi di Napoli Federico II, Dipartimento di Strutture per l'Ingegneria e l'Architettura, Via Claudio 21, 80125 Naples, Italy.
Email: georgios.baltzopoulos@unina.it

Funding information
H2020-MSCA-RISE, Grant/Award Number: 691213

Summary

Response-history nonlinear dynamic analysis is an analytical tool that often sees use in risk-oriented earthquake engineering applications. In the context of performance-based earthquake engineering, dynamic analysis serves to obtain a probabilistic description of seismic structural vulnerability. This typically involves subjecting a nonlinear numerical computer model to a set of ground-motions that represent a sample of possible realizations of base acceleration at the site of interest. The analysis results are then used to calibrate a stochastic model that describes structural response as a function of shaking intensity. The sample size of the ground-motion record set is nowadays usually governed by computation-demand constraints, yet it directly affects the uncertainty in estimation of seismic response. The present study uses analytical and numerical means to investigate the record sample size, n, required to achieve quantifiable levels of mean relative estimation error on seismic risk metrics. Regression-based cloud analysis in the context of Cornell's reliability method and incremental dynamic analysis using various intensity measures were employed to derive a relation of the form $\Delta/\sqrt{n}$, where Δ is a parameter that depends on both the dispersion of structural responses and the shape of the hazard curve at the site. For the cases examined, n can be kept in the 40 to 100 range and achieve 10% mean relative error. The study can contribute to guide engineers towards an informed a-priori assessment of the number of records needed to achieve a desired value for the coefficient of variation of the estimator of structural seismic risk.

KEYWORDS

fragility function, ground motion record selection, nonlinear dynamic analysis, seismic reliability

1 | INTRODUCTION

In the context of performance-based earthquake engineering (PBEE[1]), one of the main objectives is the probabilistic quantification of structure-specific seismic risk. A risk metric typically adopted in PBEE is the annual rate of earthquakes able to cause structural *failure*, with failure defined as the structure falling short of a seismic performance objective. This failure rate, denoted as λ_f, can be calculated via Equation 1, where $P[f|im]$ represents a function providing the probability of failure given the value of some ground motion intensity measure (IM)—typically termed the structure's *fragility function*—and λ_{im} is a measure of seismic hazard at the site, defined as the annual rate of earthquakes exceeding that value of shaking intensity.

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

The state-of-the-art for PBEE applications entails the analytical estimation of fragility functions by means of various procedures that require multiple dynamic analysis runs of a numerical model of the structure, while the evaluation of λ_{im} for various intensity levels, that is the *hazard curve*, is typically obtained by means of a probabilistic seismic hazard analysis (PSHA; eg, McGuire[2]). In principle, both PSHA and the analytical derivation of structural fragility involve the use of ground motion records; in the former case, this is implicit, as PSHA usually models IM via ground motion prediction equations (GMPE; eg, Campbell and Bozorgnia[3]), which are semi-empirical models based on recorded strong motion, while in the latter case it is explicit, as one has to select a certain number of accelerograms to conduct the analyses. The number of records typically used for nonlinear dynamic analysis is mainly dictated by the large computation times required for running complex structural models at high nonlinearity levels. However, the number of records directly determines the structural response sample size to be used in estimating fragility and, ultimately, the failure rate. As highlighted, for example, in Iervolino,[4] since these descriptors of seismic fragility and risk are inferred from samples, they are only estimates of the corresponding *true* values and are therefore affected by estimation uncertainty. In fact, the estimator of λ_f, denoted as $\widehat{\lambda}_f$, obtained using a specific sample of ground motions of certain size, can be considered at the same time a random variable and a function of the sample. If one were to re-compute $\widehat{\lambda}_f$ a number of times using different sets of accelerograms, equal in number to the first one and equivalent in characteristics to the analyst, one would keep obtaining different values for the estimator due to the *record-to-record variability* of inelastic structural response (eg, Shome et al[5]).* In order to illustrate this concept, an example will be provided in the context of incremental dynamic analysis (IDA,[8]), which is one of the simpler methods that can be used for the derivation of fragility. IDA consists of running a series of analyses for a nonlinear structure, using a suite of accelerograms, n in number, that are scaled in amplitude in order to represent a broad range of IM levels. At each IM level, a measure of structural response is registered, generically named an *engineering demand parameter* or EDP. At the conclusion of the analysis, an EDP-IM relationship is obtained, termed an IDA curve (Figure 1). At this point, it is assumed that a certain threshold value of the EDP can be defined, edp_f, such that the condition $EDP > edp_f$ can be held to signify failure (eg, violation of some *limit state*). This implies that seismic fragility can be expressed as the probability of $EDP > edp_f$ conditional to $IM = im$; ie, $P[f|im] = P[EDP > edp_f|IM = im]$. To derive the fragility curve in this way is known as the *EDP-based* approach, and although it should be noted that, in some cases, response can approach numerical instability, meaning that lack of convergence is observed so that the EDP cannot be measured, yet fragility can still be fitted.[9]

An alternative, widely used way of expressing seismic fragility is provided by the so-called *IM-based* approach. IM-based fragility entails the introduction of a new random variable, IM_f, that can be regarded as the seismic intensity able to cause structural failure.[10] By this definition, fragility can be expressed as the complementary cumulative distribution function of IM_f; ie, $P[f|im] \equiv P[IM_f \leq im]$. Returning to the IDA example, one can obtain the lowest IM value that causes each record to reach the failure criterion, by finding the height $im_{f,i}$, $i = \{1, 2, ..., n\}$, where the i-th IDA curve intersects the vertical line $EDP = edp_f$, as shown in the figure. These $im_{f,i}$ values can be considered as a sample of IM_f. Common statistical methods (Baker[11]) can then be employed to fit a parametric probability distribution model to the sample, such as the lognormal model indicated by a dark line in Figure 1B; alternatively, one may even assume that the observed sample values approximate the fragility in a non-parametric way, also depicted in Figure 1B as a stepwise function. However, if one were to repeat the procedure over a large number of times, each time performing IDA with a different set of records, it is to be expected that each repetition will lead to a different fragility curve. This variability will then reflect on the evaluated failure rate via Equation 1, and a quantitative measure can be obtained according to Equation 2:

$$CoV_{\widehat{\lambda}_f} = \frac{\sqrt{VAR\left[\widehat{\lambda}_f\right]}}{E\left[\widehat{\lambda}_f\right]}, \tag{2}$$

*There may be more sources contributing to estimation uncertainty in the fragility and the risk estimate; eg, when accounting for variability in the numerical model properties that may be based on samples of experimental data (eg, Lignos and Krawinkler[6] and Panagiotakos and Fardis[7]). Nevertheless, the present study deals exclusively with estimation uncertainty related to record-to-record variability of structural response to earthquakes, since the intended focus is on the choice of record sample size.

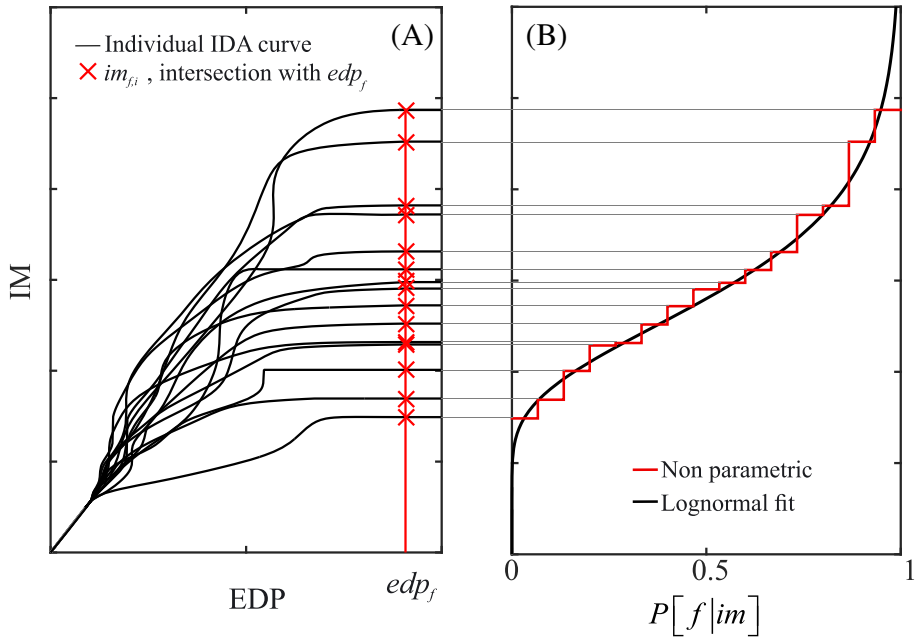

FIGURE 1 Schematic representation of possible derivation of seismic fragility functions using IDA. The intersection of the n IDA curves with the threshold EDP value defining failure A, provides a sample of failure intensities used to define a cumulative probability function B, [Colour figure can be viewed at wileyonlinelibrary.com]

where the notation $CoV_{\widehat{\lambda}_f}$ indicates the *coefficient of variation* of the failure rate estimator, and $VAR\left[\widehat{\lambda}_f\right]$, $E\left[\widehat{\lambda}_f\right]$ denote its variance and expected value, respectively.[†] In fact, in the case of unbiased estimators, $CoV_{\widehat{\lambda}_f}$ provides the mean relative (with respect to the true value) estimation error.

Past PBEE-oriented research on the topic has seen discussion on the number of records that ought be used for estimating the distribution of EDPs at a single level (or *stripe*) of intensity.[5,12] In Gehl et al,[13] the estimation uncertainty in the case of parametric fragility for simplified structural systems was examined, while Eads et al[14] examined the effect of record sample size on the confidence intervals of various fractiles of collapse fragility. The effect of estimation uncertainty on the parametric fragility model stemming from cloud analysis was considered in Jalayer et al.[15] The number of records issue has also been studied in a somewhat different context: Hancock et al[16] studied its effect under various proposals for spectral-matching and scaling of accelerograms and Reyes and Kalkan[17] paid particular attention to evaluating the ASCE/SEI-7 record selection and scaling procedure. In some of these previous studies that examined estimation uncertainty for the fragility parameters (eg, Gehl et al and Eads et al[13,14]), it was recognized that further research was needed to investigate the effect of this uncertainty to the actual end result; ie, the seismic risk estimate.

The objective of the present study is to further investigate this issue and illustrate methodologies to be used as tools for making informed decisions about the number of accelerograms to employ in earthquake engineering applications, in order to achieve a given value of $CoV_{\widehat{\lambda}_f}$. Two distinct paths are followed in tandem to reach this aim: an analytical approach based on the Cornell reliability method[18] and a computational approach based on IDA. In the former case, some additional simplifications are explored, intended to render a closed-form expression for $CoV_{\widehat{\lambda}_f}$ tractable. The analytical result is then evaluated with the aid of cloud analysis[10] performed for some simple yielding oscillators, exposed to varying hazard scenarios. In the computationally oriented case, a relatively large pool of records is assembled and used to run IDA for an assortment of low-rise code-conforming frames and simple inelastic structures. In this context, various limit states and alternative IMs are considered. Based on these IDA results, a Monte-Carlo methodology is used to simulate statistics of the estimator $\widehat{\lambda}_f$, for a wide range of record sample sizes n.

What follows is structured in such a way that the analytical treatment to the problem is presented first, starting from the classical approximate seismic reliability formulation, that is the Cornell reliability method. Following that, risk assessments using IM-based parametric and non-parametric fragility are examined. Subsequently, the procedures for

[†]These are typically unknown; however, estimates are obtainable via methods such as those presented in Iervolino.[4]

determining the suitable record sample size, by setting thresholds of tolerable estimation uncertainty, are illustrated, also considering spectral-shape-based (advanced) ground motion IMs. Finally, some concluding discussion is provided to summarize the main findings of the study.

2 | CORNELL RELIABILITY METHOD

Cornell's seismic reliability method uses some simplifying assumptions to provide an elegant closed-form solution to estimate the annual failure rate $\widehat{\lambda}_f$.[18] The analytical expression is given below as Equation 3, where k is the (absolute value of) the hazard curve's slope, calculated in logarithmic-space at IM_C, that is the IM corresponding to the median capacity, λ_{IM_C} is the annual exceedance rate of IM_C at the site of interest, β_C is the standard deviation of the structural failure threshold, and $\widehat{b}$ and $\widehat{\beta}_D$ are the slope parameter and the standard deviation of the logarithm of EDP that come with the assumption of an EDP-IM relationship as the one expressed by Equation 4.

$$\widehat{\lambda}_f = \lambda_{IM_C} \cdot e^{\frac{1}{2}\cdot\frac{k^2}{\widehat{b}^2}\cdot\left(\widehat{\beta}_D^2+\beta_C^2\right)} \tag{3}$$

$$\log(EDP) = \widehat{a} + \widehat{b}\cdot\log(IM) + \varepsilon \tag{4}$$

The parameters $\left\{\widehat{a},\widehat{b},\widehat{\beta}_D\right\}$ can be obtained via ordinary least squares linear regression in the context of *cloud analysis* (eg, Jalayer and Cornell[10]), which is typically performed considering EDP responses to a set of unscaled accelerograms from dynamic structural analysis, with $\widehat{\beta}_D$ estimated as the standard deviation of ε; ie, the *regression residual*.[19] The hat symbol over the notation serves as a reminder that these parameters are estimates obtained from finite samples of ground-motion and hence also subject to estimation uncertainty. A graphical representation of the parameters and quantities relevant to the Cornell method is provided in Figure 2.

In Iervolino,[4] the *delta method* (eg, Oehlert[20]) was used to provide analytical expressions that allow calculating the mean and variance of $\widehat{\lambda}_f$ in the framework of Cornell's method, reported herein in Equation 5, where $COV[\cdot]$ represents the covariance operator and the derivatives are those of Equation 3. By adopting $CoV_{\widehat{\lambda}_f}$ as the measure of estimation uncertainty behind the seismic risk metric $\widehat{\lambda}_f$, these two expressions can be used to derive a single formula for its quantification, via Equation 2.

$$\begin{cases} E\left[\widehat{\lambda}_f\right] = \widehat{\lambda}_f + \frac{1}{2}\cdot VAR[\widehat{a}]\cdot\frac{\partial^2\widehat{\lambda}_f}{\partial\widehat{a}^2} + \frac{1}{2}\cdot VAR\left[\widehat{b}\right]\cdot\frac{\partial^2\widehat{\lambda}_f}{\partial\widehat{b}^2} + \frac{1}{2}\cdot VAR\left[\widehat{\beta}_D^2\right]\cdot\frac{\partial^2\widehat{\lambda}_f}{\partial\left(\widehat{\beta}_D^2\right)^2} + COV\left[\widehat{a},\widehat{b}\right]\cdot\frac{\partial^2\widehat{\lambda}_f}{\partial\widehat{a}\cdot\partial\widehat{b}} \\ VAR\left[\widehat{\lambda}_f\right] = VAR[\widehat{a}]\cdot\left(\frac{\partial\widehat{\lambda}_f}{\partial\widehat{a}}\right)^2 + VAR\left[\widehat{b}\right]\cdot\left(\frac{\partial\widehat{\lambda}_f}{\partial\widehat{b}}\right)^2 + VAR\left[\widehat{\beta}_D^2\right]\cdot\left[\frac{\partial\widehat{\lambda}_f}{\partial\left(\widehat{\beta}_D^2\right)}\right]^2 + 2\cdot COV\left[\widehat{a},\widehat{b}\right]\cdot\frac{\partial\widehat{\lambda}_f}{\partial\widehat{a}}\cdot\frac{\partial\widehat{\lambda}_f}{\partial\widehat{b}} \end{cases} \tag{5}$$

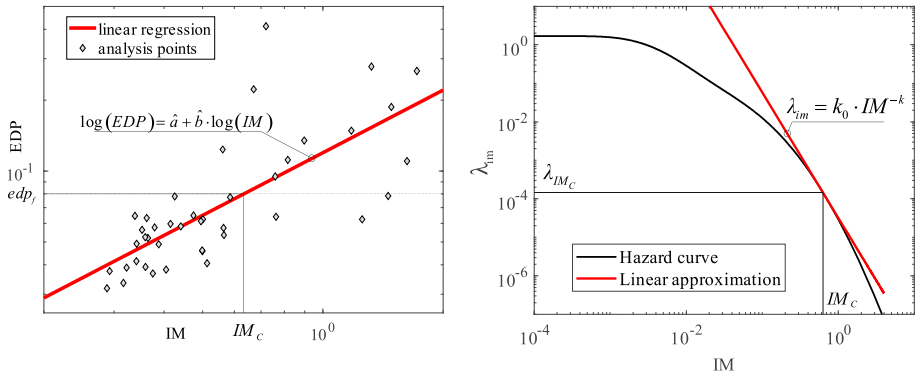

FIGURE 2 Left: Cloud analysis for a simple SDOF inelastic structure situated in Naples, Italy. Right: Local linearization of the hazard curve assumed in Cornell's seismic reliability method [Colour figure can be viewed at wileyonlinelibrary.com]

By examining applications of Cornell's method, functional to this study, it was observed that some of the terms of Equation 5 containing second derivatives of $\widehat{\lambda}_f$, and especially $\partial^2\widehat{\lambda}_f/\partial\left(\widehat{\beta}_D^2\right)^2$, were providing much smaller contributions to the summations than the other terms of the Taylor expansion. In particular, the term $VAR\left[\widehat{\beta}_D^2\right]\cdot\left[\partial\widehat{\lambda}_f/\partial\left(\widehat{\beta}_D^2\right)\right]^2$ was consistently found to be around two orders of magnitude lower than the other terms comprising $VAR\left[\widehat{\lambda}_f\right]$. Furthermore, it was observed that $E\left[\widehat{\lambda}_f\right]$ can be adequately approximated by the first two terms in the Taylor expansion; ie, $\widehat{\lambda}_f + 1/2\cdot VAR[\widehat{a}]\cdot\partial^2\widehat{\lambda}_f/\partial\widehat{a}^2$. Thus, under these hypotheses, some terms can be dropped from Equation 5 in pursuit of simplification, as they appeared less influential in practical applications (see also the applications that follow), resulting in Equation 6, which provides the approximate estimates for $E\left[\widehat{\lambda}_f\right], VAR\left[\widehat{\lambda}_f\right]$ to plug into Equation 2.

$$\begin{cases} E\left[\widehat{\lambda}_f\right] \approx \widehat{\lambda}_f\cdot\left(1+\dfrac{k^2}{2\cdot\widehat{b}^2}\cdot VAR[\widehat{a}]\right) \\ VAR\left[\widehat{\lambda}_f\right] \approx VAR[\widehat{a}]\cdot\left(\dfrac{\partial\widehat{\lambda}_f}{\partial\widehat{a}}\right)^2 + VAR\left[\widehat{b}\right]\cdot\left(\dfrac{\partial\widehat{\lambda}_f}{\partial\widehat{b}}\right)^2 + 2\cdot COV\left[\widehat{a},\widehat{b}\right]\cdot\dfrac{\partial\widehat{\lambda}_f}{\partial\widehat{a}}\cdot\dfrac{\partial\widehat{\lambda}_f}{\partial\widehat{b}} \end{cases} \tag{6}$$

With reference to Figure 2, by denoting the abscissa of the cloud analysis point corresponding to the i-th ground motion as $\log(im_i)$ and representing the sample mean and standard deviation of the logarithms of these records' IM values by $\overline{\log(im)} = 1/n\cdot\sum_{i=1}^{n}\log(im_i)$ and $s_{\log(im)} = \sqrt{1/n\cdot\sum_{i=1}^{n}\left[\log(im_i)-\overline{\log(im)}\right]^2}$, respectively, then it is known[19] that

$$\begin{cases} VAR[\widehat{a}] \approx \dfrac{\widehat{\beta}_D^2}{n}\cdot\left(1+\dfrac{\overline{\log(im)}^2}{s^2_{\log(im)}}\right) \\ VAR\left[\widehat{b}\right] \approx \dfrac{\widehat{\beta}_D^2}{n\cdot s^2_{\log(im)}} \\ COV\left[\widehat{a},\widehat{b}\right] \approx \dfrac{\widehat{\beta}_D^2}{n}\cdot\left(\dfrac{-\overline{\log(im)}}{s^2_{\log(im)}}\right) \end{cases} . \tag{7}$$

By combining Equations 2, 6, and 7, and substituting the partial derivatives of $\widehat{\lambda}_f$ (that can be obtained from Equation 3, but are also given explicitly in Iervolino[4]), then Equation 8 is obtained. In that equation, $\gamma = \log(IM_C) - \dfrac{k}{\widehat{b}^2}\cdot\left(\widehat{\beta}_D^2+\beta_C^2\right)$ and $\log(IM_C) = \left[\log\left(edp_f\right)-\widehat{a}\right]/\widehat{b}$ is the log of the median structural capacity.

$$CoV_{\widehat{\lambda}_f} \approx \frac{\widehat{\beta}_D\cdot k\cdot\sqrt{s^2_{\log(im)}+\overline{\log(im)}^2-2\cdot\overline{\log(im)}\cdot\gamma+\gamma^2}}{\sqrt{n}\cdot\widehat{b}\cdot s_{\log(im)}\cdot\left[1+\dfrac{k^2}{2\cdot\widehat{b}^2}\cdot\left(1+\dfrac{\overline{\log(im)}^2}{s^2_{\log(im)}}\right)\cdot\dfrac{\widehat{\beta}_D^2}{n}\right]} \approx \frac{\dfrac{\widehat{\beta}_D\cdot k}{\widehat{b}\cdot s_{\log(im)}}\cdot\sqrt{\overline{\log(im)^2}-2\cdot\overline{\log(im)}\cdot\gamma+\gamma^2}}{\sqrt{n}} \tag{8}$$

The simplifications, when passing unto the second (approximate) equality, are due to the observations that $s^2_{\log(im)}+\overline{\log(im)}^2 = 1/n\cdot\sum_{i=1}^{n}[\log(im_i)]^2 = \overline{\log(im)^2}$, and that for most typical earthquake engineering applications $\widehat{\beta}_D$ in natural log scale will be in the 0.20 to 0.60 range,[21] so for a number of records n of more than 10, the term $1+\left(k^2/2\cdot\widehat{b}^2\right)\cdot\left(1+\overline{\log(im)}^2/s^2_{\log(im)}\right)\cdot\left(\widehat{\beta}_D^2/n\right)$ will tend to unity. Thus, apart from the non-surprising fact that $CoV_{\widehat{\lambda}_f}$ varies inversely proportional to $\sqrt{n}$, the equation suggests that it depends, among other things, also on the shape of the hazard curve, around a region of the curve that corresponds to the structural capacity.

To better illustrate the implications of Equation 8, three yielding single-degree-of-freedom (SDOF) systems were considered; these SDOFs all had natural vibration period $T = 0.70s$ and bilinear backbone curves following a peak-oriented hysteretic rule exhibiting some mild cyclic strength degradation.[22] Each structure was assumed situated at a site of the same subsoil category (class B according to Eurocode 8[23]; ie, soil with 30-m shear wave velocity $v_{s,30}$ between 360 and 500 m/s) but characterized by varying levels of seismic hazard. Three Italian sites were chosen for this example, namely Milan (representing a low seismic hazard), Naples (medium hazard), and L'Aquila (high seismic hazard; see also Stucchi et al[24]). Hazard curves in terms of spectral acceleration at their vibration period $Sa(T = 0.7s)$ were obtained for all three sites using the REASSESS software,[25] considering the seismic source model from Meletti et al.[26] The yield displacement d_y for all structures was set corresponding to a uniform-across-sites yield annual frequency of 0.0021. The hazard curves at the three sites are shown in Figure 3. Finally, cloud analysis was performed using three sets of 42 records and the MATLAB-OpenSees[27] interface DYANAS[28] (the analysis for the SDOF situated at the Naples site is shown in Figure 2). The ground motion records were selected from the NESS flat-file[29,30] to provide coverage of a relatively wide range of $Sa(T = 0.7s)$ values.

Using all of these results, $CoV_{\widehat{\lambda}_f}$ was evaluated for various cases of failure threshold edp_f, set at ductility demands (ratio of maximum-to-yield displacement d/d_y) of four, six, and eight. The results of these calculations are presented in Table 1, along with the values assumed by the parameters involved in the intermediate computations. The corresponding example of cloud analysis for the L'Aquila SDOF structure, considering $d/d_y = 4$ as the failure threshold, is shown in Figure 3.

The last two columns of Table 1 provide a direct evaluation for the simplifications adopted in Equation 6: a comparison of the results for $CoV_{\widehat{\lambda}_f}$ given by the delta method implemented as presented in Iervolino[4]—ie, computing $CoV_{\widehat{\lambda}_f}$ by substituting Equations 5 into Equation 2—and the simplified Equation 8. It emerges, from this comparison, that the latter remains a good approximation of the former. Regarding the resulting $CoV_{\widehat{\lambda}_f}$ values themselves, it can be seen that passing from a hazard curve slope-at-median-capacity k of around 2.0 at the high-hazard site (L'Aquila) to about 3.5 for the low-hazard site (Milan), this metric of estimation uncertainty for the failure rate of nominally identical structures more than doubles. Recalling the inverse proportionality of $CoV_{\widehat{\lambda}_f}$ to $\sqrt{n}$, according to Equation 8, this result

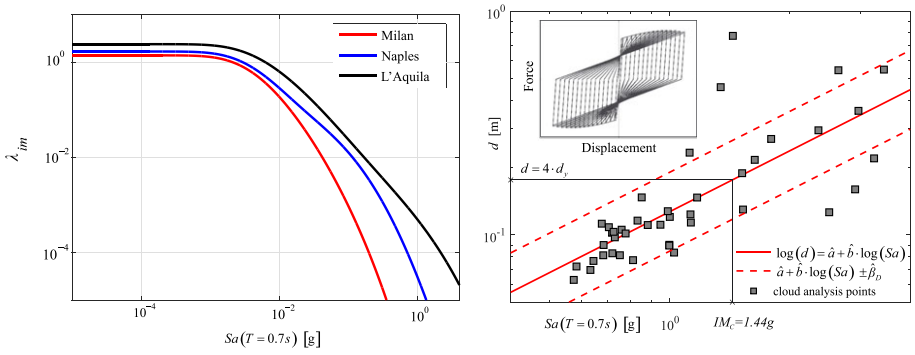

FIGURE 3 Hazard curves calculated for the three Italian sites (left) and cloud analysis for the SDOF structure assumed at a site near L'Aquila (right). The peak-oriented, degrading hysteretic behavior characterizing these inelastic SDOF oscillators used in the examples of Cornell's reliability method is also shown [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Application of Equation 8 to inelastic SDOF systems at sites with varying severity of seismic hazard

	$\frac{d}{d_y}$	$\hat{a}$	$\hat{b}$	$\hat{\beta}_D$	IM^C [g]	$k_0[\times 10^{-5}]$	k	$\lambda_{IM^C}[\times 10^{-4}]$	$\hat{\lambda}_f[\times 10^{-4}]$	$CoV_{\widehat{\lambda}_f}$ Equation 5	$CoV_{\widehat{\lambda}_f}$ Equation 8
L'Aquila	4				1.44	50.0	2.065	2.37	3.638	0.17	0.15
	6	−2.06	0.91	0.406	2.24	59.7	2.351	0.89	1.553	0.21	0.18
	8				3.08	74.1	2.569	0.41	0.795	0.28	0.25
Naples	4	−2.12	0.88	0.425	0.63	3.19	3.258	1.41	4.834	0.40	0.38
Milan	4	−2.27	0.88	0.424	0.19	0.03	3.574	1.11	4.850	0.45	0.44

implies that more than four times as many records would be needed to reduce estimation uncertainty for the risk estimate of the SDOF structure in Milan, to the same level as that of the structure in L'Aquila. This difference appears despite the fact that estimated dispersion of structural response $\widehat{\beta}_D$ and failure rate $\widehat{\lambda}_f$ are very similar for both structures. The same effect is observed at a single site, when considering the failure rate for limit states associated with progressively higher inelasticity levels: for the L'Aquila structure, changing the EDP threshold from four- to eight-times d_y resulted in an increase to $CoV_{\widehat{\lambda}_f}$ from 0.15 to 0.25, as k increased from around 2.0 to 2.5, due to IM_C moving farther to the right of the hazard curve.

The observations stemming from these illustrative examples, albeit quite evident from Equation 8, can be summarized as follows: for a fixed sample-size of ground motions, the variability, due to estimation uncertainty, of the seismic risk estimator $\widehat{\lambda}_f$ does not increase only with β_D, but also with increasing (negative) slope of the hazard curve around the intensity where mean response matches the failure threshold. In other words, the number of dynamic analyses required to reduce the estimation uncertainty of a given structure's failure rate to a specific level will increase with increasing severity of the limit state. Furthermore, for seismic risk analysis studies at sites where the curve drops off at steeper slopes, one may need a larger number of dynamic analyses to achieve a specific target coefficient of variation, with respect to a similar analysis performed at a site with a milder-sloping curve and with all else being equal structure-wise. At first sight, one might be tempted to treat these observations with some caution, as they could be influenced by the simplifying assumptions of Cornell's reliability method (see Vamvatsikos[31] for a discussion). However, the same observations are also generally confirmed under a more rigorous context of seismic risk assessment, as discussed in the following.

3 | IM-BASED FRAGILITY

In this section, the issue of estimation uncertainty is treated for cases of direct application of Equation 1, with the structure-specific fragility function $P[f|im]$ being obtained via IDA. Case-study structures used in this context are three SDOF systems at three sites, as before, and two four-story, plane, code-conforming, moment-resisting frames: a steel perimeter frame, designed to ASCE-SEI 7-05 criteria and described in the NIST GCR 10-917-8 report,[32] and a reinforced concrete bare frame (ie, without masonry infills) designed according to EN-1998-1 (see Baltzopoulos et al[33] for structural details). The only differences of the SDOF structures with respect to the previous example is that softening post-peak behavior has been modeled on their backbones to allow collapse prediction (eg, Ibarra and Krawinkler[34]) and d_y has been re-adjusted in order to ensure that, for collapse, an arbitrary yet uniform $\widehat{\lambda}_f = 4.3{\cdot}10^{-4}$ results across all sites, when $Sa(T = 0.7s)$ is used as the IM. The steel multiple-degree-of-freedom (MDOF) structure is placed at the L'Aquila site and the reinforced concrete one at the Naples site. Center-line models built in OpenSees are used for both frames; some basic geometrical information and static pushover curves are shown in Figure 4A,B. For all structural models, IDA is performed using a set of 200 ground motions that are scaled upwards until numerical instability is observed, signifying side-sway collapse. These records were selected primarily from within the NESS database (http://ness.mi.ingv.it/[29,30]) and, to a lesser extent, from the NGA-West2 database.[35] The selection criteria were to obtain records exhibiting some of the highest available naturally recorded spectral ordinates at the first-mode vibration periods of the examined structures (to keep scaling-up during IDA to a minimum), to exclude records potentially affected by near-source directivity or recorded at very soft soil sites and to avoid over-representation of any single event.

In all cases, seismic fragility functions are estimated via the 200 records according to the IM-based procedure (eg Jalayer and Cornell[10]) for the collapse limit state; additionally, limit states characterized by various threshold (edp_f) interstory drift ratios (IDRs) are considered for the two frames. This is done assuming, alternatively, lognormal and non-parametric fragility functions according to Equations 9 and 10, respectively, where $im_{f,i}$ represents the i-th record's (lowest) scaled IM value causing exceedance of the threshold (ie, referring again to Figure 1, the intersection of the i-th record's IDA curve with the vertical $EDP = edp_f$ line), $\widehat{\eta}_{IM_f}$ and $\widehat{\beta}_{IM_f}$ are the point estimates of the lognormal fragility's parameters, taken as the mean and standard deviation of the logs of $im_{f,i}$, $\Phi(\cdot)$ is the standard Gaussian function, $I_{\left(im_{f,i} \le im\right)}$ is an indicator function that returns 1 if $im_{f,i} \le im$ and 0 otherwise, and n is the total number of records.

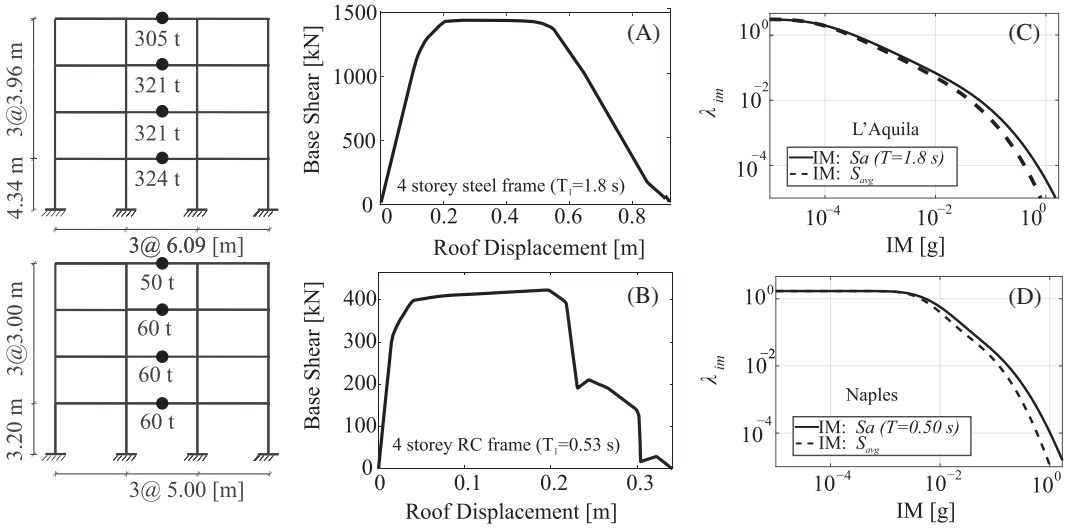

FIGURE 4 Basic information on the structures and hazard used in the examples. Dimensions and static pushover curves (first-mode proportional lateral load) for the two code-conforming, four-story, inelastic MDOF frame structures. Steel perimeter moment-resisting frame A, and reinforced concrete moment-resisting frame B. Hazard curves in terms of two IMs, $Sa(T_1)$ and S_{avg}, are shown for the L'Aquila C, and Naples site D

$$\begin{cases} P[f|im] = P\left[IM_f \le im\right] = \Phi\left[\dfrac{\log(im) - \widehat{\eta}_{IM_f}}{\widehat{\beta}_{IM_f}}\right] \\ \widehat{\eta}_{IM_f} = \dfrac{1}{n}\cdot\sum_{i=1}^{n}\log\left(im_{f,i}\right) \\ \widehat{\beta}_{IM_f} = \sqrt{\dfrac{1}{n-1}\cdot\sum_{i=1}^{n}\left[\log\left(im_{f,i}\right) - \widehat{\eta}_{IM_f}\right]^2}. \end{cases} \tag{9}$$

$$P[f|im] = P\left[IM_f \le im\right] = \frac{1}{n}\cdot\sum_{i=1}^{n} I_{\left(im_{f,i}\le im\right)} \tag{10}$$

IDA curves are initially obtained in terms of 5% damped, first-mode spectral acceleration $Sa(T_1)$ but, thanks to the use of the *hunt-and-fill* algorithm,[36] they are subsequently converted into another two more *efficient* IMs that account for spectral shape: average spectral acceleration S_{avg} ([37,38]) and I_{N_p},[39] given by Equations 11 and 12, respectively, where n_T is the number of periods considered in S_{avg} (discussion on IM efficiency to follow).

$$S_{avg} = \sqrt[n_T]{\prod_{i=1}^{n_T}\left[Sa(T_i)\right]} \tag{11}$$

$$I_{Np} = Sa(T_1)\cdot\left[S_{avg}/Sa(T_1)\right]^{0.40} \tag{12}$$

Spectral ordinates at different periods T_i are used to define these IMs for each structure. Both S_{avg} and I_{Np} for the SDOF structures use $T_i = \{0.7\text{s}, 1.0\text{s}, 1.4\text{s}\}$, while S_{avg} for the MDOF frames uses each structure's first-mode period T_1 and another three periods approximately corresponding to $\{0.3\cdot T_1, 1.5\cdot T_1, 2\cdot T_1\}$. Hazard curves for these structure-specific IMs are provided in Figure 4C,D for the frames and Figure 5 for the SDOF systems.

For each structure, limit state, and IM, the 200-record IDA-based fragilities are treated as the reference—*true*—fragility functions for the purposes of the study. Following the derivation of the fragility functions from the vectors of $im_{f,i}$ provided by IDA, Monte Carlo simulation is used to calculate the relationship of $CoV_{\widehat{\lambda}_f}$ against n. The simulation entails randomly sampling l times from the reference distribution of IM levels causing failure, IM_f, for different sample

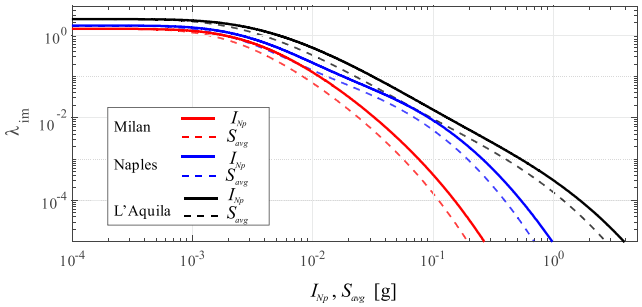

FIGURE 5 Hazard curves for the structure-specific IMs I_{Np} and S_{avg} that were used for the collapse risk estimates of the example SDOF structures at the three Italian sites of Milan, Naples, and L'Aquila [Colour figure can be viewed at wileyonlinelibrary.com]

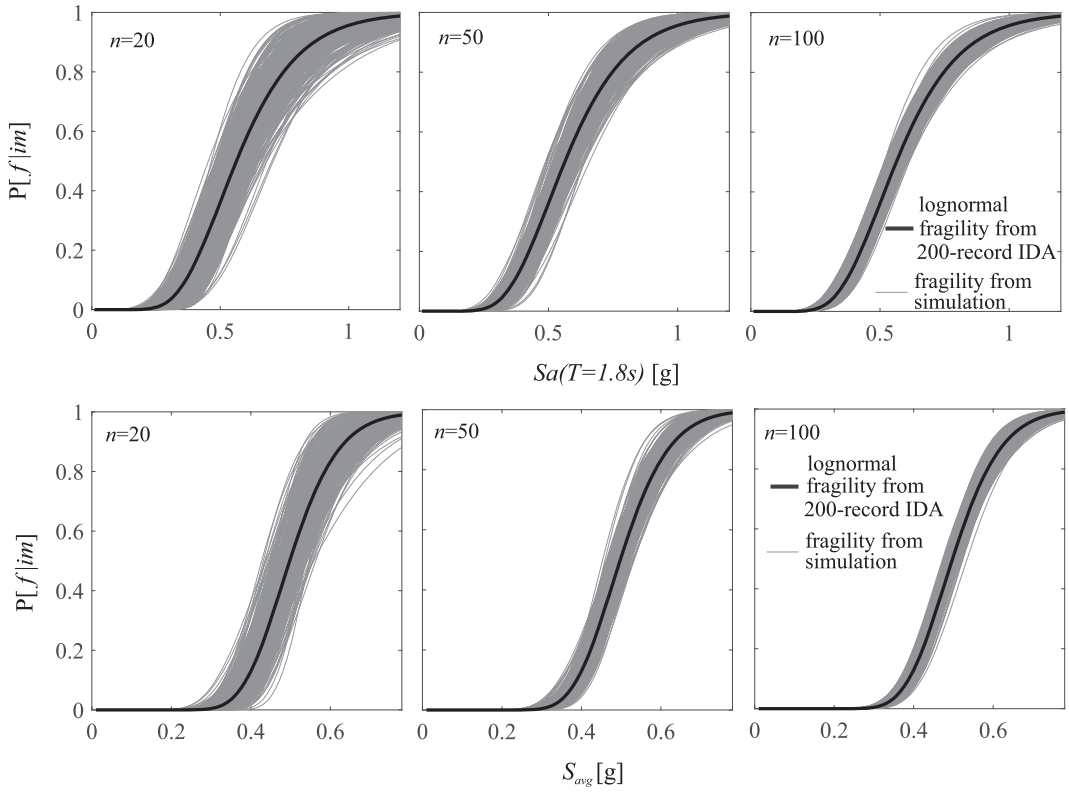

FIGURE 6 Plots of lognormal collapse fragility functions, produced during Monte-Carlo simulations that sample $n = \{20, 50, 100\}$ failure intensities from the lognormal distribution derived from the 200-record IDA of the steel, four-story frame. Top row shows fragilities in terms of $Sa(T = 1.8s)$ and bottom row in terms of S_{avg}. Each panel displays 500 simulations

sizes $n = \{2, 3, ..., 200\}$.[‡] The reference distribution sampled during the simulation is alternatively considered the lognormal function estimated from the 200-record IDA via Equation 9 and the non-parametric version of the same. In the former case, new lognormal fragility parameters are fitted to each individual sample according to Equation 9, while in the latter, Equation 10 is directly applied to the sample and considered as fragility function. Examples of the resulting simulated fragility realizations are shown in Figures 6 and 7 for the lognormal and non-parametric case, respectively. Both figures refer to the collapse fragility of the four-story steel frame at the L'Aquila site, expressed in terms of both $Sa(T = 1.8s)$ and S_{avg}. Each panel displays the reference 200-record fragility function and 500 simulated fragility realizations for sample sizes $n = \{20, 50, 100\}$. These plots provide a visual representation of the effect of estimation uncertainty on structural fragility, as increasing the sample size of structural responses results in the simulated curves clustering more tightly around the sampled reference fragility.

[‡]Although this type of simulation is reminiscent of resampling schemes such as the bootstrap, the two methods are only coincident when samples of equal size to the original 200 are being extracted.

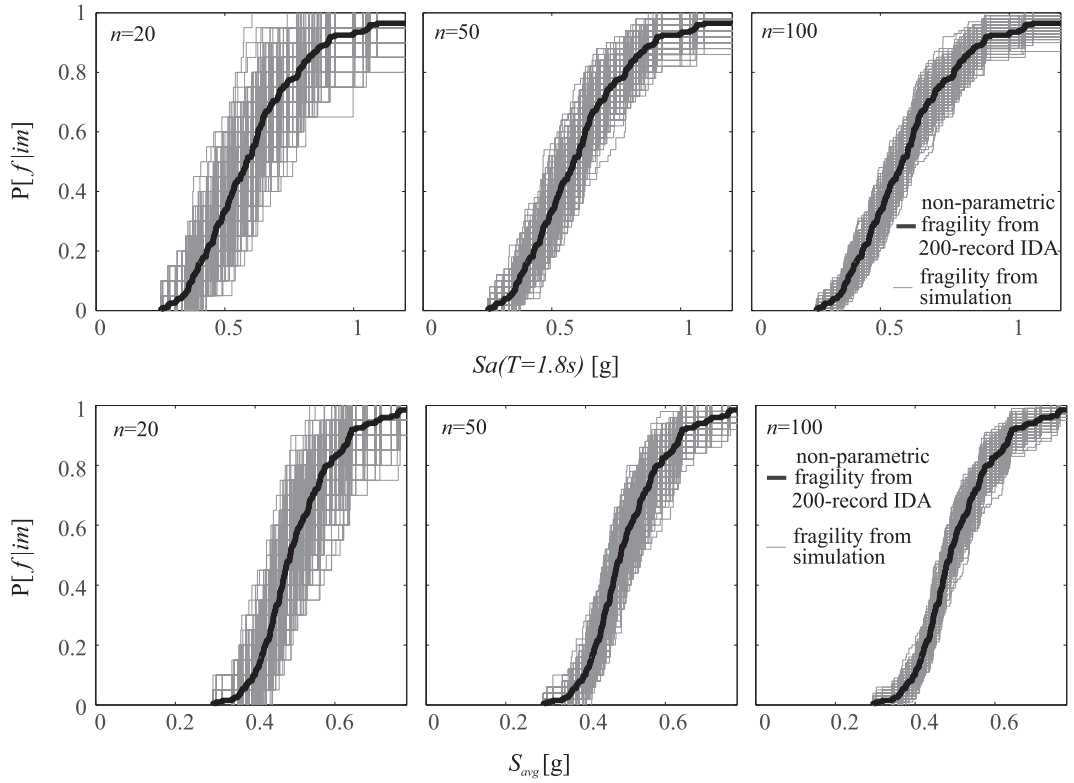

FIGURE 7 Non-parametric collapse fragility functions, produced during Monte-Carlo simulations that sample $n = \{20, 50, 100\}$ failure intensities from their empirical distribution resulting from the 200-record IDA of the four-story steel frame. Top row shows fragilities in terms of $Sa(T = 1.8s)$ and bottom row in terms of S_{avg}, with each panel displaying 500 simulations. Corresponding simulations performed under the assumption of a lognormal fragility model are shown in Figure 6

In either case, using Equation 1 leads to a point estimate of the failure rate at the j-th simulation $\widehat{\lambda}_{f,j}$, $j = \{1, 2, ..., l\}$, with a total of $l = 5000$ simulations used for each application. Finally, by substituting $VAR\left[\widehat{\lambda}_f\right]$ and $E\left[\widehat{\lambda}_f\right]$ in Equation 2 with their estimates from the Monte Carlo simulation-generated sample, $CoV_{\widehat{\lambda}_f}$ is approximated for each n according to Equation 13.

$$CoV_{\widehat{\lambda}_f} \approx \frac{\sqrt{\frac{1}{l-1}\cdot\sum_{j=1}^{l}\left(\widehat{\lambda}_{f,\mathrm{j}}-\frac{1}{l}\cdot\sum_{k=1}^{l}\widehat{\lambda}_{f,\mathrm{k}}\right)^2}}{\frac{1}{l}\cdot\sum_{j=1}^{l}\widehat{\lambda}_{f,\mathrm{j}}} \tag{13}$$

The results of this procedure are shown in Figure 8 for the three SDOF systems, in Figure 9 for the steel frame, and in Figure 10 for the reinforced concrete frame. In these figures, $CoV_{\widehat{\lambda}_f}$ is plotted against n for the various limit states and IMs considered, and the record sample sizes corresponding to coefficients of variation of 0.10 and 0.20 are denoted for select cases to facilitate comparison.

From these results, it can be observed that $CoV_{\widehat{\lambda}_f}$ from Monte Carlo simulation is very nearly inversely proportional to $\sqrt{n}$, especially for $n \geq 3$, as suggested by Equation 8. Thus, all of these curves are almost linear in the logarithms with a slope of $-1/2$; ie, $\log\left(CoV_{\widehat{\lambda}_f}\right) \approx \log(\Delta) - 1/2\cdot\log(n)$. It is, therefore, useful to calculate the log-space intercept of these curves, $\log(\Delta)$, equal to $\Delta = \sqrt{n}\cdot CoV_{\widehat{\lambda}_f}$ in linear scale, by least-squares fitting. This information is reported in Table 2 along with the logarithmic standard deviation $\left(\beta_{IM_f}\right)$ of the fragility curve for each case, where the hat from β_{IM_f} is omitted to indicate the reference 200 record-based estimate, presumed close to the true value. It is evident that

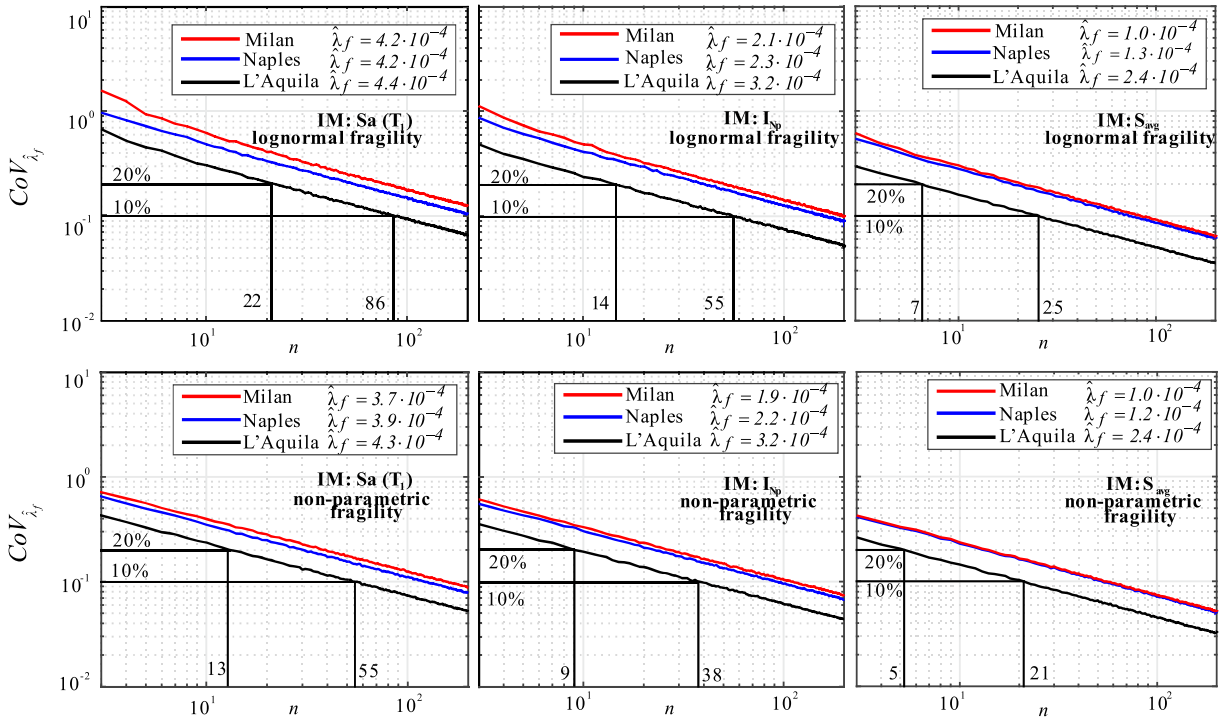

FIGURE 8 $CoV_{\widehat{\lambda}_f}$ against n calculated via Monte Carlo simulation for the three SDOF structures considered [Colour figure can be viewed at wileyonlinelibrary.com]

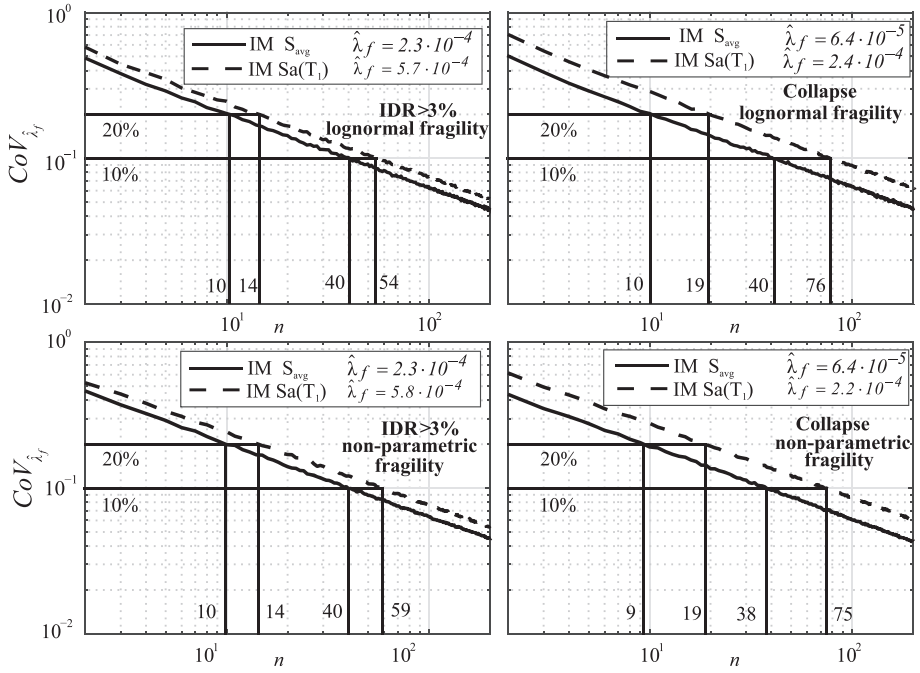

FIGURE 9 $CoV_{\widehat{\lambda}_f}$ against n calculated via Monte Carlo simulation for the four-story steel frame considered at the L'Aquila (high hazard) site

Δ bears direct correspondence to the numerator of Equation 8 and that cases exhibiting higher Δ values are subject to larger estimation uncertainty of $\widehat{\lambda}_f$ than cases with lower Δ, given parity of record sample size.

From the figures and table, it can be observed that for the limit states associated with more severe levels of inelastic response, the lognormal assumption for fragility leads to higher $CoV_{\widehat{\lambda}_f}$ than the non-parametric approach. On the other

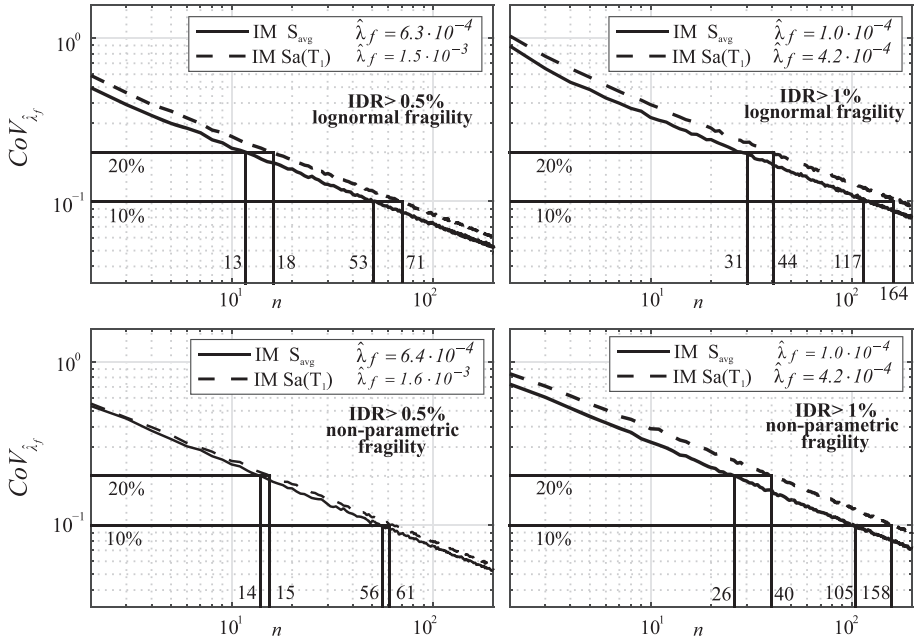

FIGURE 10 $CoV_{\widehat{\lambda}_f}$ against n calculated via Monte Carlo simulation for the four-story reinforced concrete frame considered at the Naples (medium hazard) site

hand, for limit states corresponding to smaller drifts, both the lognormal and non-parametric approach lead to similar levels of estimation uncertainty for the failure rate of the structures examined. This difference hints at the unsurprising fact that the effect of the hazard curve's shape on the failure rate's estimation uncertainty is, in reality, somewhat more complex than just due to the slope at median capacity, since both of these representations of the fragility function share the same median IM_f.

This notwithstanding, the behavior of $CoV_{\widehat{\lambda}_f}$ with varying conditions of site-specific seismic hazard observed within these results, is generally consistent with the observations made previously, on the basis of the analytical approximation derived from Cornell's method. In fact, at sites with hazard curves that slope downwards at steeper angles near the median capacity $\widehat{\eta}_{IM_f}$, the dispersion of the estimator $\widehat{\lambda}_f$ increases for structures with equal sample sizes and ostensibly equal λ_f. This is observed among the SDOF systems examined, as well as between the two MDOF structures. The same effect is also observed at the same site and structure, when different limit states are concerned: the collapse and $IDR > 3\%$ limit states for the steel frame exhibit almost identical $\widehat{\beta}_{IM_f}$ values, but the collapse case is associated with larger dispersion of $\widehat{\lambda}_f$. This can be explained by the fact that collapse is associated with a higher $\widehat{\eta}_{IM_f}$ than $IDR > 3\%$ (in terms of $Sa(T = 1.8s)$, $e^{\widehat{\eta}_{IM_f}} = 0.56g$, and 0.37g, respectively). Thus, $CoV_{\widehat{\lambda}_f}$ for the collapse limit state is influenced by a steeper portion of the hazard curve than for $IDR > 3\%$. This is analogous to what was observed when implementing Equation 8 in Cornell's method, with increasing failure thresholds for the same structure.

3.1 | Number of records and efficient intensity measures

The term *efficiency* is used in the literature to denote the property of an IM to produce lower dispersion of structural responses, conditional to that IM, than alternative, less efficient IMs (see for example Shome et al and Luco and Cornell[5,40]). Efficiency is specific to a given structural typology, EDP, and level of nonlinearity (eg, Kazantzi and Vamvatsikos[37] and Bojórquez and Iervolino[39]). Past research has shown that for EDPs related to story- and roof-drifts and for limit states nearing side-sway collapse, scalar IMs that reflect spectral shape at multiple periods, such as S_{avg} and I_{Np}, are more efficient than the classical $Sa(T_1)$.[37-41] High efficiency is typically cited as an important and desirable characteristic, precisely because it implies that a smaller number of dynamic analysis runs will be required to achieve a given dispersion level for the estimator of a seismic-risk-related parameter (eg, Cornell[12]). In fact, even though most studies that have investigated IM efficiency tend to focus on the record-to-record variability of structural response, this

TABLE 2 Logarithmic intercept Δ of $CoV_{\widehat{\lambda_f}}$ as a function of n, $\log\left(CoV_{\widehat{\lambda_f}}\right) = \log(\Delta) - 1/2 \cdot \log(n)$, calculated via 200 record IDA for various structures, IMs, site seismic hazard, and limit states

Site	Structure	Limit State	IM	β_{IM_f}	Fragility	Δ
Milan (low hazard)	Inelastic SDOF $T = 0.70$ s	Collapse	$Sa(T_1)$	0.443	Lognormal	1.874
					Non-parametric	1.247
			I_{Np}	0.340	Lognormal	1.453
					Non-parametric	1.047
			S_{avg}	0.222	Lognormal	0.945
					Non-parametric	0.745
Naples (medium hazard)	Inelastic SDOF $T = 0.70$ s	Collapse	$Sa(T_1)$	0.444	Lognormal	1.521
					Non-parametric	1.108
			I_{Np}	0.341	Lognormal	1.271
					Non-parametric	0.960
			S_{avg}	0.222	Lognormal	0.881
					Non-parametric	0.727
	Four-story reinforced concrete frame $T_1 = 0.53$ s	$IDR > 0.5\%$	$Sa(T_1)$	0.314	Lognormal	0.856
					Non-parametric	0.791
			S_{avg}	0.216	Lognormal	0.734
					Non-parametric	0.747
		$IDR > 1\%$	$Sa(T_1)$	0.389	Lognormal	1.344
					Non-parametric	1.247
			S_{avg}	0.255	Lognormal	1.113
					Non-parametric	1.020
L'Aquila (high hazard)	Inelastic SDOF $T = 0.70$ s	Collapse	$Sa(T_1)$	0.443	Lognormal	0.956
					Non-parametric	0.737
			I_{Np}	0.340	Lognormal	0.747
					Non-parametric	0.614
			S_{avg}	0.222	Lognormal	0.505
					Non-parametric	0.455
	Four-story steel moment resisting frame $T_1 = 1.82$ s	Collapse	$Sa(T_1)$	0.335	Lognormal	0.884
					Non-parametric	0.866
			S_{avg}	0.194	Lognormal	0.637
					Non-parametric	0.616
		$IDR > 3\%$	$Sa(T_1)$	0.329	Lognormal	0.745
					Non-parametric	0.761
			S_{avg}	0.227	Lognormal	0.631
					Non-parametric	0.640

is actually intended as a proxy for the estimation uncertainty underlying the risk metrics, whose reduction is the end objective. It is, therefore, quite natural to directly observe the effect of this property on estimation uncertainty, via quantities such as $CoV_{\widehat{\lambda_f}}$.

In this respect, there are two main observations to be made, on the basis of the mean relative error results obtained for the example applications. The first observation regards cases with $\Delta > 1$; ie, combinations of the more severe limit states, low-to-medium hazard, and adoption of the classical $Sa(T_1)$ as IM. In these situations, the reciprocal relation of $CoV_{\widehat{\lambda_f}}$ with $\sqrt{n}$ means that achieving an arbitrary low coefficient of variation of, say, 10% would require sample sizes in excess of 100 records, which verges on the impracticable. This hints at a more pressing need to adopt efficient IMs in such cases.

The second observation is that when switching to more efficient IMs, the reductions in the dispersion $\widehat{\beta}_{IM_f}$ are not always consistently reflected in the reduction of the number of records required to maintain any given $CoV_{\widehat{\lambda_f}}$ level. For example, switching from $Sa(T_1)$ to S_{avg} results in about 30% reduction in $\widehat{\beta}_{IM_f}$ for both the $IDR > 0.5\%$ limit state of the Naples frame and the $IDR > 3\%$ limit state of the L'Aquila frame; however, this translates to a 25% reduction in the number of records required to maintain any $CoV_{\widehat{\lambda_f}}$ level in the latter case, but only 8% in the former. Similarly,

performing the same operation for the collapse case of the SDOFs at the Milan and Naples sites results in the same 50% reduction in $\widehat{\beta}_{IM_f}$, which translates in a 75% reduction in the corresponding number of records for the structure in Milan, but only 66% for the one in Naples.

Apart from these observations, a comment should be made about the fact that the 200-record point estimate $\widehat{\lambda}_f$ shifts when switching IM. This effect is mainly related to the sensitivity of response to seismological parameters when records are scaled (see, for example, Luco and Cornell,[40]), which can be different for each IM-EDP combination. However, this is not an issue that is directly related to estimation uncertainty, and for this reason, further discussion thereof falls beyond the scope of this article.

3.2 | Target mean relative error and interquantile range of the risk estimate

The preceding discussion and illustrative applications showcased the influence of various parameters on $CoV_{\widehat{\lambda}_f}$ and provided some typical ranges of values that it is expected to take for given sample sizes. It was highlighted that the mean relative estimation error, in the case of IM-based fragility, varies with n according to $CoV_{\widehat{\lambda}_f} = \Delta/\sqrt{n}$, with Δ depending on the choice of IM and consequent record-to-record variability of structural response and on the shape of the corresponding hazard curve. The examples provided indicate that, for the simple inelastic structures examined, $\Delta \in (0.45, 1.9)$, which translates into $CoV_{\widehat{\lambda}_f}$ ranging roughly from $0.45/\sqrt{n}$ for a combination of efficient IM and high seismicity site (mild slope of the hazard curve around average structural capacity) to about $1.9/\sqrt{n}$ for the other extreme of less efficient IM and low seismicity site (steeper curve). Given that the hazard curve should be known prior to embarking on dynamic analysis and that the literature is rife with studies on the efficiency of various IMs for specific EDPs, one should be able to assess the required number of records to be $n = (\Delta/CoV_{tar})^2 \approx [(0.45 \div 1.9)/CoV_{tar}]^2$, where CoV_{tar} is a target value for the coefficient of variation of the risk estimate. Once the dynamic analysis has been concluded and structural response results have become available, it will also be possible to use other tools for assessing estimation uncertainty, such as those presented in Iervolino.[4]

In this context, it may be useful to illustrate the consequences of setting a target $CoV_{\widehat{\lambda}_f}$ of, say, 0.10 or 0.20 on the precision of the risk estimate, by linking these values to another measure of the estimator's scatter around the mean: the 5th to 95th interquantile range, $\widehat{\lambda}_{f,0.95} - \widehat{\lambda}_{f,0.05}$.[42] This statistic can be calculated from the approximations to the distribution of $\widehat{\lambda}_f$ obtained by the Monte Carlo simulations for each n, an example of which is provided in Figure 11. The examples in the figure refer to the *IDR > 1%* limit state for the reinforced concrete frame and collapse of the steel frame, in both cases using S_{avg} as IM, and are given for $n = \{20, 50, 100\}$. The distribution plots clearly showcase the reduction in dispersion of the risk estimator, with increasing sample size of records.

On the other hand, the interquantile range $\widehat{\lambda}_{f,0.95} - \widehat{\lambda}_{f,0.05}$ is plotted in Figure 12, normalized by the point estimate at $n = 200$: $\widehat{\lambda}_{f^{(n=200)}}$. This is shown for two cases of collapse annual rate: that of the steel four-story frame, using non-parametric fragility in terms of S_{avg}, and that of the SDOF structure at the Naples site, using non-parametric fragility in terms of $Sa(T_1)$ this time. The underlying implication behind plotting the ratio $\widehat{\lambda}_f/\widehat{\lambda}_{f^{(n=200)}}$ is that $\widehat{\lambda}_{f^{(n=200)}}$ should be close to the true rate, and thus the interquantile range appears as a percentage of that value. From the figure, it can be observed that normalized $\widehat{\lambda}_{f,0.95} - \widehat{\lambda}_{f,0.05}$ ranges corresponding to the same value of $CoV_{\widehat{\lambda}_f}$ are almost identical between the two cases, even if the latter requires more records to achieve those $CoV_{\widehat{\lambda}_f}$ values than the former. Furthermore, it can be seen that the width of this range shrinks from around 0.70 at $CoV_{\widehat{\lambda}_f} = 0.20$, to about 0.33 at $CoV_{\widehat{\lambda}_f} = 0.10$. These normalized $\widehat{\lambda}_{f,0.95} - \widehat{\lambda}_{f,0.05}$ widths suggest that mean relative errors in the environs of 10%, and not larger than 20%, should be regarded as acceptable target levels of accuracy in estimation, as higher $CoV_{\widehat{\lambda}_f}$ levels imply that the width of this interquantile range would approach the reference $\widehat{\lambda}_{f^{(n=200)}}$ value in size. In terms of the required number of records, setting $CoV_{tar} = 0.10$ and assuming that, as discussed previously, Δ will be maintained under 1.0 by judicious selection of IM to avoid impractical sample sizes, n would result within a range from 40 to 100 ground motions, depending on IM efficiency with respect to all relevant EDPs and site-specific hazard and

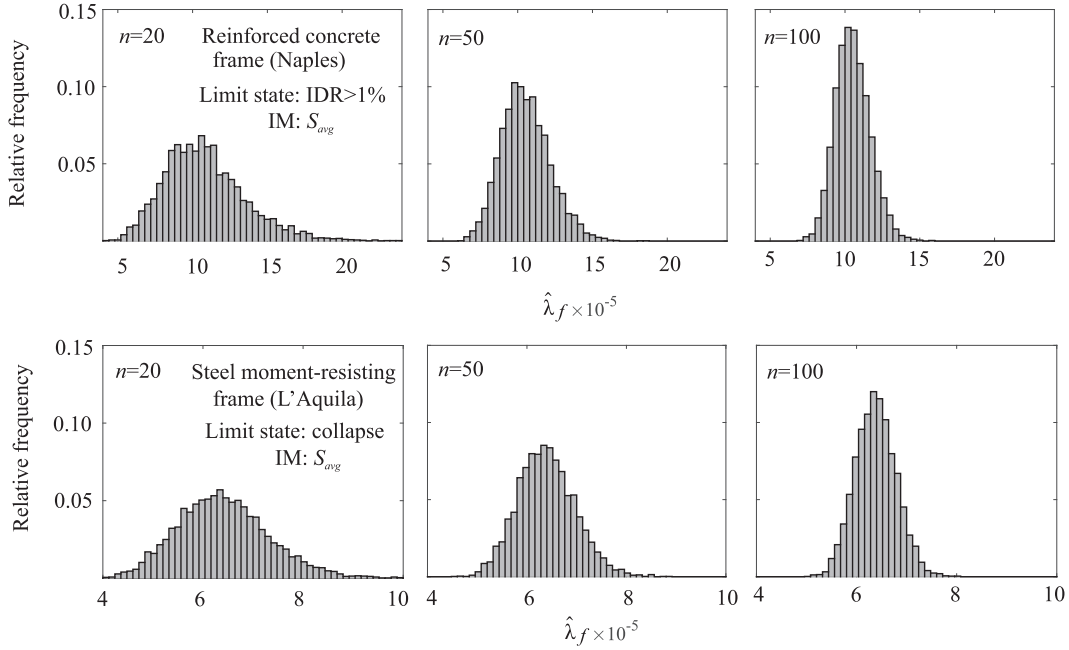

FIGURE 11 Distributions of the seismic risk estimators $\widehat{\lambda}_f$, approximated via Monte Carlo simulation, for record sample sizes $n = \{20, 50, 100\}$. The top row corresponds to the *IDR > 1%* limit state of the reinforced concrete frame situated at the Naples site and the bottom row to collapse of the steel frame situated at the L'Aquila site. The IM employed in both cases is S_{avg}. The ordinate reports relative frequency; ie, the number of simulated $\widehat{\lambda}_f$ point estimates contained in each bin of the histogram divided by the total number of simulations

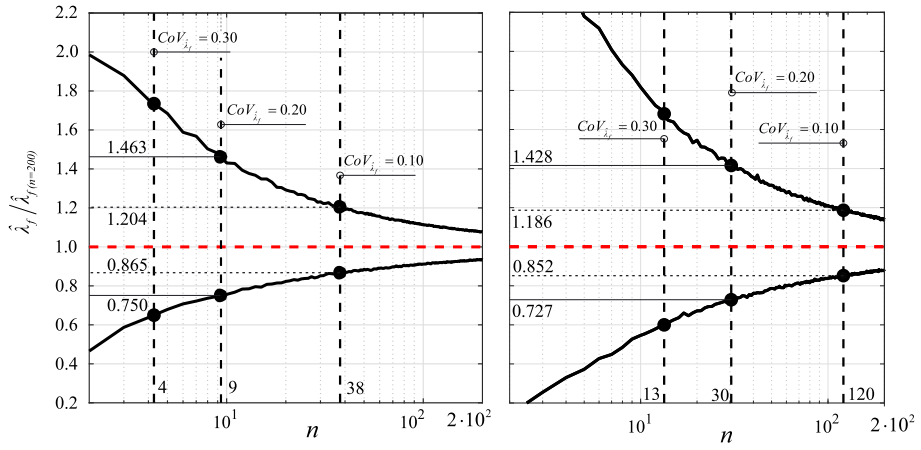

FIGURE 12 Interquantile range $\widehat{\lambda}_{f,0.95} - \widehat{\lambda}_{f,0.05}$ as a function of n, normalized by the 200-record point estimate $\widehat{\lambda}_{f^{(n=200)}}$. The $\widehat{\lambda}_{f,0.95} - \widehat{\lambda}_{f,0.05}$ range was calculated from the Monte Carlo approximations of the distributions of $\widehat{\lambda}_f$ for various sample sizes: Cases of collapse limit state for the steel frame (non-parametric fragility in terms of S_{avg}, left panel) and collapse limit state for the SDOF system located at the Naples site (non-parametric fragility in terms of $Sa(T_1)$, right panel). Record numbers corresponding to $CoV_{\widehat{\lambda}_f}$ of 0.10, 0.20, and 0.30 for each case are highlighted by dashed black lines [Colour figure can be viewed at wileyonlinelibrary.com]

considering the results of the MDOF frames as representative. In fact, for the simple inelastic SDOF structures, this number starts from as low as around 20 ground motions.

4 | DISCUSSION AND CONCLUSIONS

The complexity of numerical models that simulate the seismic response of structures in the nonlinear range grows at a pace that rivals advances in computing power. This leads to computational costs for analytical seismic risk assessments

that remain at constantly high levels, rendering the issue of the appropriate number of response-history analyses to run, ever topical. The present article advocates the use of quantitative criteria to determine the sample size of ground motion records, based on the statistical inference concept of estimation uncertainty. This offers earthquake engineers a means of making an informed decision, by weighing computation costs against precision of the risk estimates. The study focused exclusively on the use of naturally recorded acceleration time-histories (possibly modified by scaling-in-amplitude only), which is becoming the norm, in part thanks to the widespread recent availability of online strong motion repositories. In order to quantify estimation uncertainty, the coefficient of variation of the estimator of annual failure rate, $CoV_{\widehat{\lambda}_f}$, was chosen. The main conclusions to be drawn from this study are listed here.

- A closer look at Cornell's simplified, analytical seismic reliability formulation, showed that the record-to-record dispersion of structural responses is only part of the story, with $CoV_{\widehat{\lambda}_f}$ also depending on the shape of the site-specific hazard curve. The implication of this finding is that for structures located at sites characterized by different seismicity levels, different numbers of runs may be required to reach the same level of confidence in the risk estimate between sites, even for the case of similar structures expected to exhibit the same failure rate. The effect of hazard on $CoV_{\widehat{\lambda}_f}$ can be summarized by the slope of the curve near intensity levels that are most relevant for causing failure: the steeper the local drop-off of the curve the greater the estimation uncertainty behind $\widehat{\lambda}_f$. These observations were generally corroborated by numerical simulations, based on IDA of simple inelastic systems and two code-conforming frames.
- Both analytical and numerical investigations showed that the mean relative error of the failure rate follows a relation of the type $CoV_{\widehat{\lambda}_f} = \Delta/\sqrt{n}$ and suggest that the parameter Δ, as a rule of thumb (based on the applications developed in this study), varies between around 0.45 and 1.9 for various limit states that can be defined in terms of maximum drift thresholds, ranging from moderate inelasticity to side-sway collapse. The lower value corresponds to cases where advanced, efficient scalar seismic IMs are employed combined with a mild-sloping hazard curve with slope $k \approx 2$ around median structural capacity, and the higher one to the use of traditional IMs, such as first-mode spectral acceleration, coupled with locally steeper hazard curves, eg, with slope $k \approx 3.5$.
- It was observed that, for nominally equivalent structures, variations in the choice of IM and/or site-specific hazard can cause the number of records required to achieve a level of $CoV_{\widehat{\lambda}_f}$ around 0.10 to vary from the low tens to a couple of hundred. This, once again, highlights the importance of selecting efficient IMs in PBEE, which becomes an almost necessity in the case of collapse failure estimation at low-seismicity areas, keeping the number of ground motions needed to achieve a 10% coefficient of variation for the failure rate estimate within the range of 40 to 100.
- Finally, it was observed that the reduction in the dispersion of structural response via efficient IMs, despite its importance, does not tell the whole tale with regard to the corresponding reduction in computational costs, which can also be site dependent.

On a concluding note, in the applications presented herein, the general trend was that less records are needed to reach the same level of mean relative error in the risk estimates for sites exposed to higher seismic hazard, than for sites exposed to lower hazard levels; however, this result may be a rule-of-thumb that is not necessarily generalizable independently of the shape of particular site-specific hazard curves. Generally speaking, it was shown that the results and observations presented can be useful in practical applications of seismic risk assessment, by providing a quantitative basis for determining the required number of records for risk-targeted dynamic analysis. This number can be defined by setting a target of desired precision for the risk estimate in terms of the mean relative error, $CoV_{\widehat{\lambda}_f}$.

ACKNOWLEDGEMENTS

The study presented in this paper was developed within the activities of ReLUIS (*Rete dei Laboratori Universitari di Ingegneria Sismica*) for the project ReLUIS-DPC 2014 to 2018, as well as within the H2020-MSCA-RISE-2015 research project EXCHANGE-Risk (Grant Agreement Number 691213). Professor Dimitrios Vamvatsikos (National Technical University of Athens, Greece) is gratefully acknowledged for providing one of the numerical models used in this study, as well as professor Sinan Akkar (Bogazici University, Turkey) for his constructive review.

ORCID

Georgios Baltzopoulos https://orcid.org/0000-0002-0460-6558
Iunio Iervolino https://orcid.org/0000-0002-4076-2718

REFERENCES

1. Cornell CA, Krawinkler H. Progress and challenges in seismic performance assessment. *PEER Cent News*. 2000;3(2):1-4.
2. McGuire RK. Probabilistic seismic hazard analysis and design earthquakes: closing the loop. *Bull Seismol Soc Am*. 1995;85(5):1275-1284. https://doi.org/10.1016/0148-9062(96)83355-9
3. Campbell KW, Bozorgnia Y. NGA-West2 ground motion model for the average horizontal components of PGA, PGV, and 5% damped linear acceleration response spectra. *Earthq Spectra*. 2014;30(3):1087-1114. https://doi.org/10.1193/062913EQS175M
4. Iervolino I. Assessing uncertainty in estimation of seismic response for PBEE. *Earthq Eng Struct Dyn*. 2017;46(10):1711-1723. https://doi.org/10.1002/eqe.2883
5. Shome N, Cornell CA, Bazzurro P, Carballo JE. Earthquakes, records, and nonlinear responses. *Earthq Spectra*. 1998;14(3):469-500. https://doi.org/10.1193/1.1586011
6. Lignos DG, Krawinkler H. Development and utilization of structural component databases for performance-based earthquake engineering. *J Struct Eng*. 2013;139(8):1382-1394. https://doi.org/10.1061/(ASCE)ST.1943-541X.0000646
7. Panagiotakos TB, Fardis MN. Deformation of reinforced concrete at yielding and ultimate. *ACI Struct J*. 2001;98(2):135-147. https://doi.org/10.14359/10181
8. Vamvatsikos D, Cornell CA. Incremental dynamic analysis. *Earthq Eng Struct Dyn*. 2002;31(3):491-514. https://doi.org/10.1002/eqe.141
9. Shome N, Cornell CA. Structural seismic demand analysis: consideration of collapse. *8th ACSE Specialty Conference on Probabilistic Mechanics and Structural Reliability* 2000(3): PMC2000–119.
10. Jalayer F, Cornell CA. Direct probabilistic seismic analysis: implementing non-linear dynamic assessments. Stanford University, 2003.
11. Baker JW. Efficient analytical fragility function fitting using dynamic structural analysis. *Earthq Spectra*. 2015;31(1):579-599. https://doi.org/10.1193/021113EQS025M
12. Cornell C. Hazard, ground motions and probabilistic assessments for PBSD. In: Fajfar P, Krawinkler H, eds. *International Workshop on Performance-Based Seismic Design Concepts and Implementation*. Stanford, California: Blume Center; 2004.
13. Gehl P, Douglas J, Seyedi DM. Influence of the number of dynamic analyses on the accuracy of structural response estimates. *Earthq Spectra*. 2015;31(1):97-113. https://doi.org/10.1193/102912EQS320M
14. Eads L, Miranda E, Krawinkler H, Lignos DG. An efficient method for estimating the collapse risk of structures in seismic regions. *Earthq Eng Struct Dyn*. 2013;42(1):25-41. https://doi.org/10.1002/eqe.2191
15. Jalayer F, De Risi R, Manfredi G. Bayesian cloud analysis: efficient structural fragility assessment using linear regression. *Bull Earthq Eng*. 2015;13(4):1183-1203. https://doi.org/10.1007/s10518-014-9692-z
16. Hancock J, Bommer JJ, Stafford PJ. Numbers of scaled and matched accelerograms required for inelastic dynamic analyses. *Earthq Eng Struct Dyn*. 2008;37(14):1585-1607. https://doi.org/10.1002/eqe.827
17. Reyes JC, Kalkan E. How many records should be used in an ASCE/SEI-7 ground motion scaling procedure? *Earthq Spectra*. 2012;28(3):1223-1242. https://doi.org/10.1193/1.4000066
18. Cornell CA, Jalayer F, Hamburger RO, Foutch DA. Probabilistic basis for 2000 SAC federal emergency management agency steel moment frame guidelines. *J Struct Eng*. 2002;128(4):526-533. https://doi.org/10.1061/(ASCE)0733-9445(2002)128:4(526)
19. Draper NR, Smith H. *Applied Regression Analysis*. 3rd ed. New York: John Wiley & Sons; 1998.
20. Oehlert GW. A note on the delta method. *Am Stat*. 1992;46(1):27-29. https://doi.org/10.1080/00031305.1992.10475842
21. FEMA. FEMA-350: recommended seismic design criteria for new steel moment-frame buildings. SAC Joint Venture: 2000.
22. Lignos DG, Krawinkler H. Deterioration modeling of steel components in support of collapse prediction of steel moment frames under earthquake loading. *J Struct Eng*. 2011;137(11):1291-1302. https://doi.org/10.1061/(ASCE)ST.1943-541X.0000376
23. CEN. *EN 1998-1:* Eurocode 8—design of structures for earthquake resistance. Part 1: General rules, seismic actions and rules for buildings. European Committee for Standardization; 2004.
24. Stucchi M, Meletti C, Montaldo V, Crowley H, Calvi GM, Boschi E. Seismic hazard assessment (2003-2009) for the Italian building code. *Bull Seismol Soc Am*. 2011;101(4):1885-1911. https://doi.org/10.1785/0120100130
25. Chioccarelli E, Cito P, Iervolino I, Giorgio M. REASSESS V2.0: software for single- and multi-site probabilistic seismic hazard analysis. *Bull Earthq Eng*. 2018. https://doi.org/10.1007/s10518-018-00531-x
26. Meletti C, Galadini F, Valensise G, et al. A seismic source zone model for the seismic hazard assessment of the Italian territory. *Tectonophysics*. 2008;450(1–4):85-108. https://doi.org/10.1016/j.tecto.2008.01.003

27. McKenna F. OpenSees: a framework for earthquake engineering simulation. *Comput Sci Eng*. 2011;13(4):58-66. https://doi.org/10.1109/MCSE.2011.66

28. Baltzopoulos G, Baraschino R, Iervolino I, Vamvatsikos D. Dynamic analysis of single-degree-of-freedom systems (DYANAS): a graphical user interface for OpenSees. *Eng Struct*. 2018;177:395-408.

29. Pacor F, Felicetta C, Lanzano G, et al. NESS v1.0: a worldwide collection of strong-motion data to investigate near source effects. *Seismol Res Lett*. 2018;89(6):2299-2313.

30. Luzi L, Puglia R, Russo E, et al. The engineering strong-motion database: a platform to access pan-European accelerometric data. *Seismol Res Lett*. 2016;87(4):987-997. https://doi.org/10.1785/0220150278

31. Vamvatsikos D. Accurate application and second-order improvement of SAC/FEMA probabilistic formats for seismic performance assessment. *J Struct Eng*. 2014;140(2):04013058. https://doi.org/10.1061/(ASCE)ST.1943-541X.0000774

32. NEHRP Consultants Joint Venture. Evaluation of the FEMA P-695 methodology for quantification of building seismic performance factors (NIST GCR 10–917-8). 2010.

33. Baltzopoulos G, Chioccarelli E, Iervolino I. The displacement coefficient method in near-source conditions. *Earthq Eng Struct Dyn*. 2015;44(7):1015-1033. https://doi.org/10.1002/eqe.2497

34. Ibarra LF, Krawinkler H. Global collapse of frame structures under seismic excitations. PEER Report, 2005.

35. Bozorgnia Y, Abrahamson NA, Al Atik L, et al. NGA-West2 research project. *Earthq Spectra*. 2014;30(3):973-987. https://doi.org/10.1193/072113EQS209M

36. Vamvatsikos D, Cornell CA. Applied incremental dynamic analysis. *Earthq Spectra*. 2004;20(2):523-553. https://doi.org/10.1193/1.1737737

37. Kazantzi AK, Vamvatsikos D. Intensity measure selection for vulnerability studies of building classes. *Earthq Eng Struct Dyn*. 2015;44(15):2677-2694. https://doi.org/10.1002/eqe.2603

38. Eads L, Miranda E, Lignos DG. Average spectral acceleration as an intensity measure for collapse risk assessment. *Earthq Eng Struct Dyn*. 2015;44(12):2057-2073. https://doi.org/10.1002/eqe.2575

39. Bojórquez E, Iervolino I. Spectral shape proxies and nonlinear structural response. *Soil Dyn Earthq Eng*. 2011;31(7):996-1008. https://doi.org/10.1016/j.soildyn.2011.03.006

40. Luco N, Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthq Spectra*. 2007;23(2):357-392. https://doi.org/10.1193/1.2723158

41. Tsantaki S, Adam C, Ibarra LF. Intensity measures that reduce collapse capacity dispersion of P-delta vulnerable simple systems. *Bull Earthq Eng*. 2017;15(3):1085-1109. https://doi.org/10.1007/s10518-016-9994-4

42. Mood AM, Graybill FA, Boes DC. *Introduction to the Theory of Statistics*. 3rd ed. New York: McGraw-Hill; 1974.

How to cite this article: Baltzopoulos G, Baraschino R, Iervolino I. On the number of records for structural risk estimation in PBEE. *Earthquake Engng Struct Dyn*. 2019;48:489–506. https://doi.org/10.1002/eqe.3145