Non-parametric fragility curves for bridges using recorded ground motions

C. V. Mai[1], B. Sudret[1], K.R. Mackie[2], B. Stojadinovic[3], K. Konakli[1]
[1] Chair of Risk, Safety & Uncertainty Quantification, Institute of Structural Engineering, ETH Zürich, Stefano-Franscini-Platz 5, CH-8093 Zürich, Switzerland
[2]Department of Civil, Environmental and Construction Engineering, University of Central Florida, 4000 Central Florida Blvd. , Orlando FL 32816-2450, USA
[3]Chair of Structural Dynamics & Earthquake Engineering, Institute of Structural Engineering, ETH Zürich, Stefano-Franscini-Platz 5, CH-8093 Zürich, Switzerland
email: mai@ibk.baug.ethz.ch, sudret@ibk.baug.ethz.ch, kevin.mackie@ucf.edu, stojadinovic@ibk.baug.ethz.ch, konakli@ibk.baug.ethz.ch

ABSTRACT: Seismic fragility curves are commonly used to assess the vulnerability of structures to earthquakes by representing the probability of structural failure as a function of an earthquake intensity measure. The classical approach for computing fragility curves assumes the curves to have a lognormal shape. This approach is therefore parametric. Since fragility curves play an important role in the framework of performance-based earthquake engineering, it is of utmost importance either to validate the accuracy of the lognormal assumption or to propose an assumption-free approach to compute these curves. In a recent work, the authors validated the lognormal assumption for a linear steel frame. However, in a more realistic case study of this steel frame with non-linear material behavior, the lognormal assumption showed insufficient accuracy. In this paper, we consider a different type of structure, *i.e.* a typical reinforced-concrete bridge column subject to recorded ground motions. We compute fragility curves for the column, first with the classical parametric approach and subsequently with a non-parametric method based on kernel density estimation. Two different intensity measures are considered, namely the peak ground acceleration and the pseudo-spectral acceleration. The results show the limitations of the classical lognormal approach for this type of structure and prove kernel density estimation to be a promising tool for establishing seismic fragility curves.

KEYWORDS: non-parametric fragility curves, concrete bridge columns, lognormal assumption, kernel density estimation

1 INTRODUCTION

In the framework of performance-based earthquake engineering, seismic fragility curves are a common tool for evaluating the vulnerability of structures to earthquakes. The vulnerability is represented by the probability (or the likelihood) that the considered structure fails to fulfill a prescribed safety criterion (concerning stresses, strains, displacements, *etc.*). The failure probability is a function of the earthquake severity. The latter is usually described by quantitative intensity measures such as the peak ground acceleration PGA and the pseudo-spectral acceleration at the fundamental period T_1 of the structure $Psa(T_1)$ (see *e.g.* [1] for a wide variety of intensity measures).

Fragility curves can be computed analytically as follows. One first represents the considered structure by an analytical model (usually using a finite element code). Analyses of the model subject to a set of earthquake ground motions are conducted from which structural responses are extracted. Fragility curves are built using the collected results together with the known ground motion intensity. Note that the uncertainty (both aleatory and epistemic) represented by the fragility comes from multiple sources. In this approach, we are assuming that the only source is randomness in the ground motion.

As one can see in the definition, fragility curves have no prescribed shape. However, these curves are usually computed under the assumption that they have the shape of a lognormal cumulative distribution function. In a recent paper, [2] validated this assumption in the case of a linear steel frame structure subject to a large set of synthetic ground motions. However, the lognormal curves built for a similar structure with non-linear behavior showed some discrepancies with the reference curves [3].

In this paper, we consider a different type of structure namely a reinforced concrete (RC) bridge column under recorded ground motions. We compute the fragility curves of the RC column using the lognormal approach and a non-parametric approach based on the kernel density estimation technique [3]. Subsequently, we compare the curves obtained with the different approaches.

The paper is organised as follows: in Section 2, the model of a concrete bridge column in the finite element code OpenSees [4] is presented. In Section 3, we present the set of recorded ground motions that was selected for the time-history analyses in Section 4. The lognormal and kernel density estimation-based approaches are described in Section 5, followed by the results and discussions in Section 6.

2 REINFORCED CONCRETE BRIDGE COLUMN AND PERFORMANCE LEVELS

2.1 *Reinforced concrete bridge column*

Typical California highway overpass bridges are considered. More precisely, we are interested in bridges with single-column bents where the column has a uniform circular cross section over the entire column height above the ground and continues into an integral pile shaft foundation (see Figure 1).

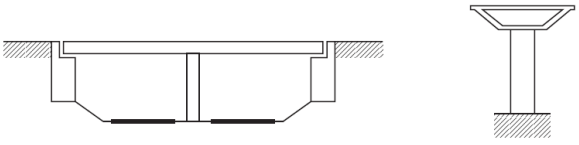

Figure 1: Bridge longitudinal and transverse configurations [5]

Structural damages of all the major bridge components (*e.g.* decks, columns, abutments, bearings, *etc.*) may be observed after earthquake events. Among multiple failure modes of bridges, the column failure is of particular interest, since it can lead to the malfunction or even the collapse of a bridge, see *e.g.* [6, 7].

In this paper, the circular reinforced concrete (RC) bridge column of diameter $D = 1.2$ m and height $H = 6$ m is modelled

by a 3D cantilever element (Figure 2). The weight of the superstructure (*e.g.* deck, girders, non-structural elements, *etc.*) is represented by a concentrated vertical load $N = 300$ kN. The column has perimeter longitudinal and transverse spiral reinforcement: 30 longitudinal reinforcing bars (32 mm diameter, 819 mm^2 area each) are evenly placed at a radius defined by a 50 mm concrete cover and the diameter of the transverse reinforcement. The spiral transverse rebars (19 mm diameter, 284 mm^2 area each) are used to confine the concrete and contribute to the shear resistance of the section during earthquakes. The spacing between transverse reinforcement is 0.1 m.

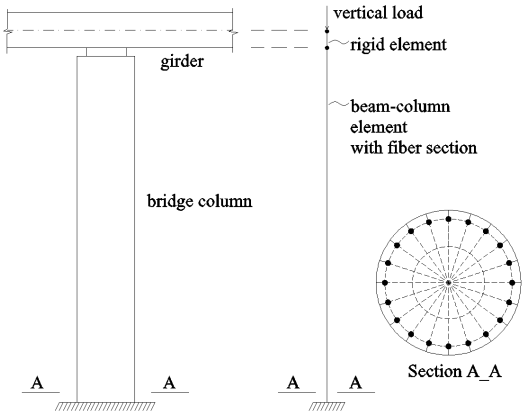

Figure 2: Bridge column model

The column is modelled in the finite element code OpenSees as a fiberized nonlinear beam-column element. The use of fiber section allows modelling of the distribution of plasticity along the column height and over the cross section. 60 and 24 layers were used to model the core and cover concrete respectively. The concrete material behavior is based on the Kent-Scott-Park stress-strain relation [8] (material `Concrete02` in OpenSees). The maximum compressive stress of unconfined concrete is $f_c = 45$ MPa. The characteristics of confined concrete are calculated according to the Mander model [9]. The crushing strain is set equal to 2.5%. The longitudinal reinforcing bars are represented by a layer of fibers at their respective positions. An elastic-plastic trilinear model was used for the steel material. This is basically the additive model of three `UniaxialMaterial` models in OpenSees, namely one elastic model, one elastic perfectly-plastic model and one kinematic harderning model with added Giuffré-Menegotto-Pinto transitions on the loading and unloading loops. The steel has an initial stiffness $E_s = 200,000$ MPa and a post-yield stiffness equal to $0.015E_s$ ($3,000$ MPa). The yield stress is $f_y = 462$ MPa corresponding to a yield strain 0.24%. The steel continues to hardern after yielding until a strain of 8% and then it softens at a slope equal to 7.5% of the initial stiffness up to the ultimate strain of 14%. If the tensile (resp. compressive) strain falls above (resp. below) 14%, the material is assumed to be in failure. The softening behavior models the strength degradation of the material. Moreover, this steel model allows one to account for the Bauschinger effect. The nonlinear geometry effects ($P - \Delta$ effects) are taken into consideration as well by using the `LinearWithPDelta` geometric transformation. The nonlinear shear and torsion behaviors are calculated and aggregated in the section. The reader is referred to [5] for more details on the modelling of the concrete column in OpenSees.

2.2 *Performance levels of concrete bridge columns*

During an earthquake, the reinforced concrete columns might experience the following damages sequentially [10, 11]: concrete cracking, yield of the longitudinal reinforcement, concrete cover spalling, fracture of the transverse reinforcement, buckling and fracture of the longitudinal rebars, concrete core crushing, shear failure and axial load failure. Damages are classified into five levels, namely no damage, minor, moderate, major damage and local failure/collapse. [12] provides photos of damaged columns from laboratory experiments and past earthquakes that illustrate the different damage levels.

The initiation and progression of damages are related to the functionality of the bridge in terms of performance levels. A five-level performance evaluation approach was proposed in [11]. The performance levels range from cracking, yielding, initiation of local mechanism, full development of local mechanism and finally strength degradation. The corresponding socio-economic description of these performance levels are fully operational, operational, life safety, near collapse and collapse. These five qualitative performance levels are linked to quantitative design parameters (*e.g.* steel strain, concrete strain, plastic rotation, *etc.*) among which the *drift ratio* will be considered in this paper.

It is worth noting that there exist formulations which allow one to predict drift ratios corresponding to certain damages such as concrete cover spalling, longitudinal reinforcement buckling [13]. The predicted values of the drift ratio limits are functions of the material strength, diameter of reinforcing bars, column diameter, axial load ratio, *etc.* In this paper, for the sake of simplicity we use the drift limits for concrete bridge columns recommended by two sources [10, 14] for the operational and life safety levels. These drift limits are shown in Table 1.

Table 1: Bridge performance/design parameter

Reference	Level	Description	Damage	Drift δ_o
[10]	II	Operational	Minor	0.01
[10]	III	Life safety	Moderate	0.03
[14]	II	Operational	Minor	0.005
[14]	III	Life safety	Moderate	0.015

Note that different sources recommend different drift limits for the same performance level. The accuracy of these values, however, is not discussed herein.

3 GROUND MOTIONS SELECTION

The ground motions utilized in this paper are records from the PEER Strong Motion Database that were collected and previously used by [5]. Initially, the records are chosen in the framework of the *bin approach*, *i.e.* the records are classified in four bins according to the earthquake moment magnitude M, the closest distance to fault rupture R and the local soil type. These are characteristic of the non-near-field motions ($R > 15$ km) recorded in California. All of them have three orthogonal component accelerograms corresponding to fault-normal, fault-parallel and vertical directions. Subsequently, the ground motion set is enriched with some near-field records. These represent ground motions from high-magnitude earthquakes recorded at distances less than 15 km. Finally, a total number of 531 triplets was selected, covering both near- and far-field earthquakes (R varies from 10 to 100 km) and a relatively large range of moment magnitude (M varies from 6 to 7.6). Figure 3 depicts the distribution of the selected records in

the $M - R$ space. The intensity measures of the earthquakes (*e.g.* PGA, Psa, *etc.*) are extracted from the records using a data extraction routine. In this paper, Psa stands for $Psa(T_1)$ which is the pseudo-spectral acceleration at the fundamental period T_1 of the column.

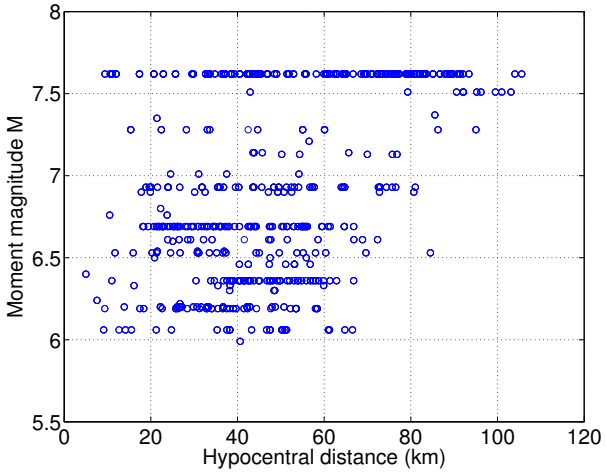

Figure 3: Distribution of selected ground motion records in the $M - R$ space

4 NON-LINEAR DYNAMIC ANALYSES

We first run modal analysis of the bridge column which results in the initial elastic period $T_1 = 0.535$ s in both the longitudinal and transverse directions. To demonstrate the force-deformation relation of the column, we conduct full cycles of loading/ unloading behavior where the column top displacement varies between 0 m, 0.1 m, -0.1 m, 0.15 m, -0.15 m and 0 m. Then the column is reset to its initial condition and pushed over monotonically. Given the similarity in the geometry and the reinforcement in the two directions, the loading/ unloading pattern and the push-over curves are similar. In both directions, the column fails at a top displacement equal to 0.35 m, which corresponds to a drift ratio $\Delta = 0.05$. Figure 4 shows the loading/ unloading behavior of the column and the push-over curve in the transverse direction.

The column is then subject to the selected ground motions. Having at hand three orthogonal component accelerograms (two horizontal and one vertical) for each earthquake record, we could afford to run the three-dimensional time-history analyses of the column. The fault-normal component is set to hit the column in the longitudinal direction (bridge-span direction), the fault-parallel component hits the column in the transverse direction and the vertical component strikes in the up-down direction. It is worth noting that the ground motion records were not scaled during the analyses but were kept as recorded. For illustration, the top displacement of the column subject to the Northridge-01 earthquake motion recorded at the station Simi Valley - Katherine Rd ($M = 6.69$, $R = 21.32$ km) is depicted in Figure 5. One can observe that the vertical motion does not pose a critical threat to the column compared to the horizontal motions.

During the analyses, the intensity measures (IM) of the ground motions (PGA, Psa) and the corresponding maximal responses (drift ratios Δ) in the transverse direction were collected. Finally, 531 paired samples (IM_i, Δ_i) were available.

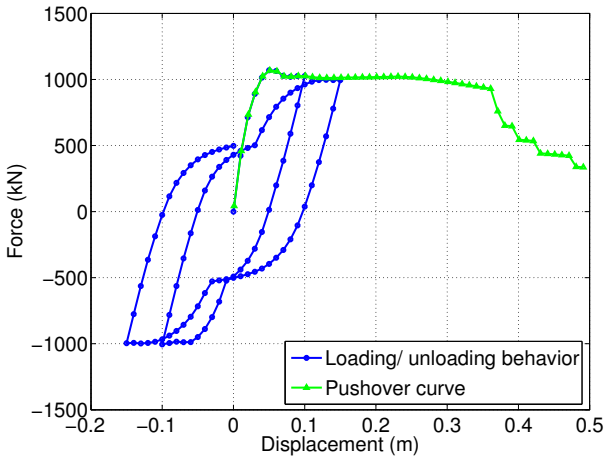

Figure 4: Loading/ unloading behavior of the column and push-over curve in the transverse direction

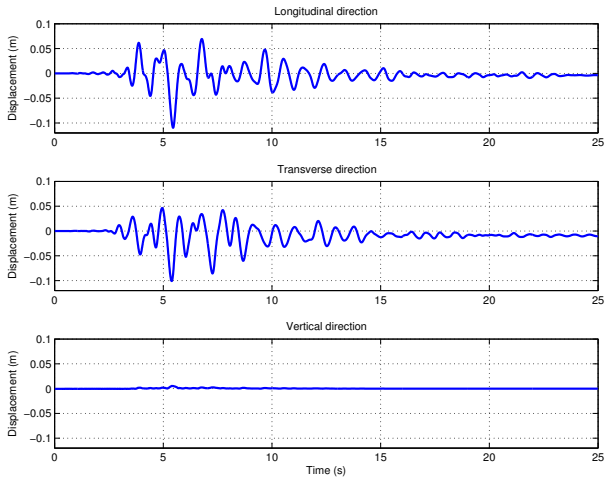

Figure 5: Top displacement of the column subject to a specific earthquake

5 SEISMIC FRAGILITY CURVES

The fragility curves, by definition, represent the failure probability of the considered structure with respect to a safety criterion (*e.g.* the drift ratio limit δ_o) given the intensity measure (IM) of the earthquake motion (*e.g.* PGA, Psa). The failure probability is expressed explicitly by the conditional probability that the drift ratio Δ attains or exceeds the threshold δ_o given IM as follows:

$$\text{Frag}(IM;\, \delta_o) = \mathbb{P}[\Delta \geq \delta_o | IM] \qquad (1)$$

In this section, we present the classical lognormal approach and a recently proposed non-parametric approach to compute fragility curves.

5.1 *Classical approach*

The approach that has been typically employed so far assumes that the fragility curves have the shape of a lognormal cumulative distribution function (CDF) as a function of the intensity measure. In recent years, some different shapes have been adopted for fragility curves, namely the normal, exponential, Weibull and Gamma CDFs [15, 16]. All these curves belong to a class of fragility curves that is considered parametric in the sense that their shape is preselected. The lognormal curves are,

however, the most commonly used. The lognormal curves can be computed using two techniques that are described next.

5.1.1 *Linear regression-based approach*

Suppose that the design parameter (*e.g.* the drift ratio Δ) is related to the ground motion IM (*e.g.* PGA, Psa) by a probabilistic seismic demand model (PSDM) which reads:

$$\log \Delta = A \, \log IM + B + \zeta \, Z; \; Z \sim \mathcal{N}(0,1) \tag{2}$$

where $\log(\cdot)$ denotes the natural logarithm. This probabilistic model assumes that Δ has a lognormal distribution whose log-mean value is a linear function of $\log IM$. As a result, the fragility curve in Eq. (1) is reformulated as follows:

$$\begin{aligned} \widehat{\text{Frag}}(IM;\, \delta_o) &= \mathbb{P}\left[\log \Delta \geq \log \delta_o\right] \\ &= \Phi\left(\frac{\log IM - (\log \delta_o - B)/A}{\zeta / A}\right) \end{aligned} \tag{3}$$

where $\Phi(t) = \int_{-\infty}^{t} e^{-u^2/2}/\sqrt{2\pi}\,\mathrm{d}u$ is the Gaussian CDF. Eq. (3) implies that the fragility curve has the shape of a lognormal CDF as a function of IM. Consider the value IM_o that satisfies $\log \delta_o = A \, \log IM_o + B$. For the sake of simplicity, hereafter we refer to the quantities $\mu = (\log \delta_o - B)/A = \log IM_o$, $\sigma = \zeta/A$ and $m = e^{\mu} = IM_o$ as "log-mean", "log-standard deviation" and "median", respectively. We underline that these terms are merely used to describe the shape of the fragility curve and do not represent properties of a certain lognormally distributed random variable. The log-standard deviation is a measure of the steepness of the fragility curve, whereas the median determines the position IM_o where the curve attains the value 0.5.

This approach relies on the computation of the parameters of the PSDM in Eq. (2). To this purpose, a linear regression technique in the logarithmic scale is usually applied. The parameters A, B, ζ are determined using least square errors estimation.

5.1.2 *Maximum likelihood estimation-based approach*

Assume that the fragility curves can be written in the following general form:

$$\widehat{\text{Frag}}(IM;\, \delta_o) = \Phi\left(\frac{\log IM - \log m}{\sigma}\right) \tag{4}$$

in which m is the median and σ is the log-standard deviation of the lognormal curves. [17] proposed the use of maximum likelihood estimation to determine these parameters. Denote by ω the event that the damage corresponding to the prescribed threshold δ_o occurs. Assume that $Y(\omega)$ is a random variable with Bernoulli distribution, *i.e.* Y takes the values 1 and 0 with probability $\text{Frag}(\cdot;\, \delta_o)$ and $1 - \text{Frag}(\cdot;\, \delta_o)$. Considering the independent Bernoulli experiments where the structure is subject to ground motions $i = 1, \ldots, N$, the likelihood function of $(m,\, \sigma)$ reads:

$$\mathcal{L} = \prod_{i=1}^{N} \left[\text{Frag}(IM_i;\, \delta_o)\right]^{y_i} \, \left[1 - \text{Frag}(IM_i;\, \delta_o)\right]^{1-y_i} \tag{5}$$

where IM_i is the intensity measure of the i^{th} motion and y_i takes the value 1 or 0 depending on whether the structure subject to the i^{th} motion has a drift greater than δ_o or not. The parameters $(m,\, \sigma)$ are determined by maximizing the likelihood function. In practice, a straightforward optimization technique is applied on the log-likelihood function, *i.e.* :

$$(m;\, \sigma) = \arg\max \log \mathcal{L} \tag{6}$$

5.2 *Non-parametric approach*

In this section, a non-parametric approach to compute fragility curves, recently proposed by [3], is presented. This approach is based on the *kernel density estimation* (KDE) technique. KDE is an evolution of the histogram technique that is commonly used for estimating the probability density function (PDF) of random variables. As defined earlier, the fragility represents the conditional probability that Δ attains or exceeds a safety threshold δ_o given IM=a. Let us rewrite Eq. (1) as:

$$\text{Frag}(a;\, \delta_o) = \int_{\delta_o}^{+\infty} f_{\Delta}(\delta | IM = a)\,\mathrm{d}\delta \tag{7}$$

By definition, the conditional PDF in Eq. (7) is given by:

$$f_{\Delta}(\delta | IM = a) = \frac{f_{\Delta, IM}(\delta, a)}{f_{IM}(a)} \tag{8}$$

in which $f_{\Delta,IM}(\cdot)$ (resp. $f_{IM}(\cdot)$) is the joint distribution of the vector $(\Delta,\, IM)$ (resp. the marginal distribution of IM). These joint and marginal PDFs will be estimated by means of KDE.

Let us start with $f_{IM}(a)$. Given a sample set $\{IM_i, \ldots, IM_N\}$ of the random variable IM, its marginal PDF is obtained as [18]:

$$\hat{f}_{IM}(a) = \frac{1}{N h_{IM}} \sum_{i=1}^{N} K\left(\frac{a - IM_i}{h_{IM}}\right) \tag{9}$$

where h_{IM} is the *bandwidth parameter* and $K(\cdot)$ is the *kernel function* which is positive and integrates to one. A standard normal PDF is usually adopted for the kernel, *i.e.* $K(x) \equiv \varphi(x) = e^{-x^2/2}/\sqrt{2\pi}$.

Similarly, given a sample set $\{(\Delta_i,\, IM_i),\, i = 1, \ldots, N\}$ one can estimate the joint PDF using the multivariate standard normal kernel as follows:

$$\begin{aligned} \hat{f}_{\Delta,IM}(\delta, a) &= \frac{1}{2\pi N \left|\boldsymbol{H}\right|^{1/2}} \times \\ &\sum_{i=1}^{N} \exp\left[-\frac{1}{2}\begin{pmatrix}\delta - \Delta_i \\ a - IM_i\end{pmatrix}^{\intercal} \boldsymbol{H}^{-1} \begin{pmatrix}\delta - \Delta_i \\ a - IM_i\end{pmatrix}\right] \end{aligned} \tag{10}$$

where $\boldsymbol{H}$ is the *bandwidth matrix* and $|\boldsymbol{H}|$ is its determinant.

The proposed estimator of the fragility curve eventually reads:

$$\widehat{\text{Frag}}(a;\,\delta_o) = \int_{\delta_o}^{+\infty} \hat{f}_{\Delta}(\delta|IM=a)\mathrm{d}\delta = \frac{h_{IM}}{2\pi\,|\boldsymbol{H}|^{1/2}} \times$$

$$\frac{\int_{\delta_o}^{+\infty} \sum_{i=1}^{N} \exp\left[-\frac{1}{2}\begin{pmatrix}\delta-\Delta_i\\ a-IM_i\end{pmatrix}^{\mathsf{T}} \boldsymbol{H}^{-1} \begin{pmatrix}\delta-\Delta_i\\ a-IM_i\end{pmatrix}\right] \mathrm{d}\delta}{\sum_{i=1}^{N} \varphi\left(\frac{a-IM_i}{h_{IM}}\right)} \tag{11}$$

The accuracy of the estimate relies mostly on the evaluation of the bandwidth parameters $\boldsymbol{H}$ and h_{IM}. A short review of the existing techniques as well as an engineering-oriented tool for computing the bandwidth parameters can be found in [19]. Note that the choice of the kernel function does not affect significantly the density estimation, hence the resulting fragility curves.

Preliminary investigations have shown that KDE-based curves obtained for pairs $(IM,\,\Delta)$ are characterized by large uncertainty for large IM values and drift limits. This is due to the scarcity of observations in the upper tail of the distribution of an IM (extreme earthquake events occur at relatively low frequencies) and also, due to the large dispersion of highly nonlinear structural responses. To reduce this effect, inspired by the functional form of the probabilistic seismic demand models, [19] proposed to apply the KDE method on the paired data in the logarithmic scale. Using the standard random variable transformation technique, [19] proved that the fragility curves in Eq. (7) can be equivalently estimated as follows:

$$\widehat{\text{Frag}}(a;\,\delta_o) = \int_{\log\delta_o}^{+\infty} \hat{f}_U(u|V=\log a)\,\mathrm{d}u \tag{12}$$

in which $U=\log\Delta$ and $V=\log IM$. In the subsequent example application, KDE-based fragility curves are developed using the logarithmic transformation in Eq. (12).

6 RESULTS AND DISCUSSION

6.1 *Fragility curves for concrete bridge column*

In this section, using the data collected from the nonlinear time-history analyses, we establish the fragility curves of the concrete bridge column by means of the lognormal and KDE-based approaches. Both linear regression (LR) and maximum likelihood estimation (MLE) techniques are used to determine the parameters of the lognormal curves. We build fragility curves corresponding to the various drift ratio limits presented in Section 2.2 ($\delta_o = 0.005,\ 0.01,\ 0.015,\ 0.03$). Two intensity measures are considered, namely the peak ground acceleration (PGA) and the pseudo-spectral acceleration (Psa), which are among the most commonly used in fragility analysis.

The non-parametric KDE-based approach is assumption-free and purely based on the distribution of the collected data. In [3, 19] we validated the accuracy of the KDE-based curves through comparisons with fragility curves obtained with a Monte Carlo simulation approach using a large number of analyses of a steel frame structure subject to synthetic ground motions. Therefore, in this paper we consider the non-parametric KDE-based curves as reference to assess the accuracy of the lognormal curves.

As described in Section 5, the lognormal approach consists in assuming that the fragility curves have the shape of a lognormal CDF and then estimating the parameters of this function. Using the LR technique, the parameters of the lognormal curves are indirectly derived by fitting a linear model to the paired data $(\log IM,\,\log\Delta)$. Figure 6 (resp. Figure 7) depicts the paired data $\{(PGA_i,\Delta_i)\}$ (resp. $\{(Psa_i,\Delta_i)\}$) in the log-scale and the fitted PSDM using LR. The coefficient of determination R^2 of the LR model considering PGA (resp. Psa) as IM is 0.729 (resp. 0.963). We observe that using Psa as IM leads to smaller dispersion *i.e.* smaller variance of the data (ζ in Eq. (2)) as compared to using PGA. This is expected since Psa is a structure-specific IM. We note in Figure 7 that for small values of Psa (Psa<2 m/s^2) the data are almost perfectly correlated. However, a larger dispersion is observed for higher values of Psa. In the MLE approach, for each drift ratio the observed failures are modeled as outcomes of a Bernoulli experiment and the fragility parameters are determined by maximizing the respective log-likelihood function. Finally, the KDE approach requires estimation of the bandwidth parameter and the bandwidth matrix. In the present example, these are determined as h= 0.2166 and $\boldsymbol{H} = \begin{bmatrix}0.1039 & 0.0781\\ 0.0781 & 0.0724\end{bmatrix}$ (resp. h= 0.2218 and $\boldsymbol{H} = \begin{bmatrix}0.0528 & 0.0501\\ 0.0501 & 0.0482\end{bmatrix}$) when PGA (resp. Psa) is considered as IM.

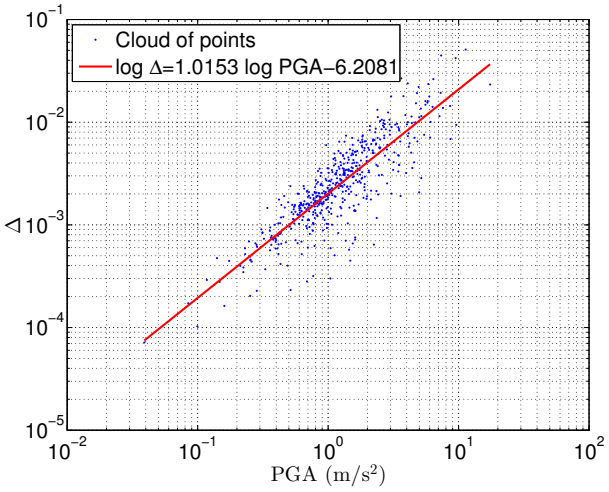

Figure 6: Paired data $\{(PGA_i,\Delta_i)\,,i=1,\,\ldots\,,531\}$ and linear regression in the logarithmic scale

For the two IMs and the three drift limits of interest, Table 2 lists the medians and log-standard deviations of the lognormal curves fitted with the LR and MLE approaches as well as the medians for the KDE-based curves determined as the points where the curves attain the value 0.5. The corresponding fragility curves are depicted in Figure 8 (resp. Figure 9) when the considered IM is PGA (resp. Psa). When PGA is used as IM (see Figure 8), for all drift limits considered, the lognormal curves obtained by LR are close to those obtained by MLE. The lognormal curves exhibit discrepancies from the KDE-based curves, which tend to be larger for larger PGA values and larger drift limits. Table 2 shows that the differences between the medians of the lognormal and the KDE-based curves are larger for the limit $\delta_o = 0.03$. When Psa is

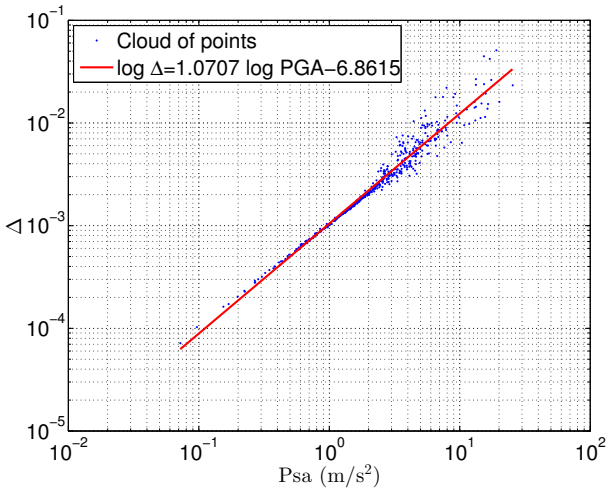

Figure 7: Paired data $\{(Psa_i, \Delta_i), i = 1, \ldots, 531\}$ and linear regression in the logarithmic scale

used as IM (see Figure 9), the MLE-based lognormal curves are in a fair agreement with the non-parametric curves, particularly for the smaller drift limits. However, the lognormal curves obtained with LR exhibit significant discrepancies from the KDE-based curves, which become larger with increasing drift limit.

Table 2: Parameters of fragility curves

IM	δ_o	Approach	IM_o (m/s^2)	σ
PGA	0.005	LR	2.4500	0.4980
		MLE	2.3763	0.4637
		KDE	2.2255	
	0.01	LR	4.8490	0.4980
		MLE	4.8696	0.4568
		KDE	4.4817	
	0.015	LR	7.2292	0.4980
		MLE	7.0111	0.4806
		KDE	6.6859	
	0.03	LR	14.3080	0.4980
		MLE	13.2471	0.4370
		KDE	11.0232	
Psa	0.005	LR	4.3055	0.1748
		MLE	4.2402	0.2162
		KDE	4.4817	
	0.01	LR	8.2256	0.1748
		MLE	8.4852	0.3209
		KDE	8.1662	
	0.015	LR	12.0122	0.1748
		MLE	12.7630	0.3800
		KDE	13.4637	
	0.03	LR	22.9489	0.1748
		MLE	23.9702	0.3688
		KDE	20.0855	

Note that the LR estimation leads to far less accurate lognormal curves for Psa than for PGA (as compared to the KDE-based curves), although the R^2 coefficient of the linear fit is higher for Psa. This is due to the fact that the assumption of *homoscedastic errors*, inherent in Eq. (2), is not valid for the specific data set (Psa, Δ), as one can observe in Figure 7. The respective errors in the linear model have close to zero variance for Psa<2 m/s^2 and increasing variance for larger values of Psa. Accordingly, by considering a constant variance over the entire range of Psa values, we underrestimate the actual variance in the regions of interest in fragility analysis. Note that by transformation from the log-scale to the the linear scale of the data, these effects appear amplified. By providing a distinct pair of median and log-standard deviation at each level of δ_o, MLE outperforms LR in the estimation of lognormal fragility curves. In Table 2, we observe that the log-standard deviations estimated by MLE are significantly larger than those obtained by LR.

The above analysis demonstrates that the accuracy of the lognormal curves strongly depends on (i) the considered drift limit, (ii) the IM used to represent earthquake severity and (iii) the technique used to estimate the parameters of the curves. Overall, the lognormal curves exhibit larger deviations from the KDE-based curves for larger drift limits. The MLE-based curves are closer to the KDE-based curves for Psa as IM than for PGA. However, for Psa as IM, the LR approach yields unacceptably inaccurate estimates of fragility. It is worth noting that different intensity measures are recommended in the literature for different types of structures [20, 21]. For instance, in establishing fragility curves for highway bridges, [20] recommend use of PGA and Psa for synthetic motions, whereas they recommend use of cumulative absolute velocity for recorded motions. For reinforced concrete frame buildings, [21] show that in terms of efficiency (*i.e.* smaller dispersion in structural response) peak ground velocity (PGV) is the optimal IM, followed by structure-specific IMs (*e.g.* Psa), whereas PGA is disqualified. However, [21] mention certain limitations of PGV for the case of low-rise RC frames. It becomes evident that the selection of the optimal among numerous different IMs (*e.g.* 65 options are listed in [1]) depends on several factors and thus, is not trivial. This highlights the superiority of the KDE approach, which allows the computation of assumption-free fragility curves that are reliable independently of the choice of the intensity measure.

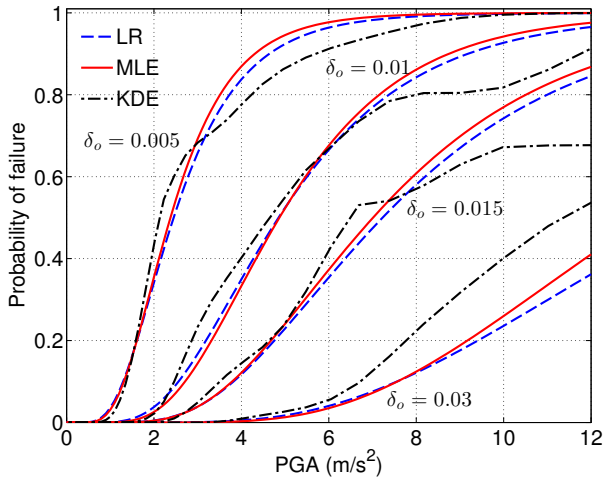

Figure 8: Fragility curves using PGA

6.2 *Statistics of fragility curves using bootstrap resampling*

Let us now examine the uncertainty in the estimated fragility curves. To this end, we employ the so-called *bootstrap resampling technique* that allows evaluation of the variability of a statistical measure for a given sample [22]. We first obtain $M = 100$ resamples of the original data set $\{(IM_i, \Delta_i), i = 1, \ldots, N\}$ by random sampling *with replacement*. These represent the so-called bootstrap samples. Each bootstrap sample has the same size as the original data

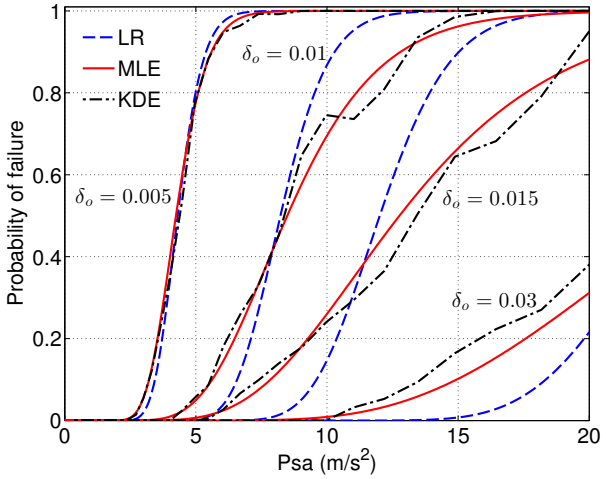

Figure 9: Fragility curves using Psa

set, but some of the original observations may appear multiple times while others may be missing. For each bootstrap sample, we compute the fragility curves using the techniques presented in Section 5. Finally, we perform statistical analysis of the M so-obtained sets of bootstrap curves and extract statistical measures of interest. In the following, this procedure is employed to estimate bootstrap medians and 95% confidence intervals for the fragility curves shown in Figure 8 and Figure 9.

Figure 10 shows bootstrap medians and confidence intervals for the fragility curves obtained with the MLE- and KDE-based approaches. As shown above, the LR approach for estimation of the fragility parameters may yield unacceptable errors and thus, the respective bootstrap statistics are not examined. In the upper (resp. lower) graphs of Figure 10, the considered IM is PGA (resp. Psa). As an example demonstration of the use of bootstrap resampling in fragility analysis, results are shown only for the two lower drift limits (*i.e.* $\delta_o = 0.005, 0.01$). Fragility curves for the higher drift limits are characterized by larger uncertainty due to a smaller number of available earthquake records that cause deformations in the respective ranges. Figure 10 shows that in all cases the bootstrap median curves are consistent with the curves obtained from the initial data set, which proves the stability and reliability of the proposed approaches. Next, we examine the uncertainty in the fragility curves in terms of the confidence intervals. For a specified IM type and drift limit, the confidence intervals become wider with increasing IM values. This is due to the smaller number of observations corresponding to these regions. For the lower values of a considered IM, confidence intervals for the MLE- and KDE-based approaches have similar widths. For larger IM values, confidence intervals for the MLE-based curves tend to be narrower due to the fact that the MLE approach arbitrarily prescribes the shape of the fragility curves. It was shown earlier that in these regions, the MLE-based curves may deviate from the KDE-based curves, which are considered to represent the "true" fragilities. We also note in Figure 10 that use of Psa as IM leads to narrower confidence intervals than use of PGA. This is due to the higher correlation of the former with structural response.

7 CONCLUSIONS

The paper proposed the kernel density estimation (KDE) approach for developing seismic fragility curves as an alternative to the classical approach that presumes the curves have the shape of a lognormal-CDF. KDE is an efficient non-parametric approach that can be used to compute fragility curves when the actual shape of these curves is not known as well as to validate or calibrate parametric fragility curves. The accuracy of the KDE approach was recently established through comparison with results from another non-parametric approach based on Monte-Carlo simulation [19].

In an example illustration, we applied the KDE approach to compute seismic fragility curves for a typical concrete bridge column using a large number of time-history analyses with recorded ground motions. Fragility curves were developed for four column drift limits considering two ground motion intensity measures, namely the peak ground acceleration (PGA) and the pseudo-spectral acceleration (Psa). The KDE-based curves were compared with fragility curves obtained with the classical parametric lognormal approach. The parameters of the latter were estimated with two techniques, namely by linear regression assuming a linear probabilistic seismic demand model and by maximum likelihood estimation.

Considering the KDE-based curves as the reference fragility curves, we showed that for the specific data, the accuracy of the lognormal curves was highly dependent on the drift limit, the method of parameter estimation and the considered intensity measure. The accuracy of both parameter estimation approaches deteriorated with increasing drift limit. Particularly large deviations from the KDE-based curves were observed for lognormal curves versus Psa with parameters estimated by means of linear regression. In the same example application, we demonstrated the use of bootstrap resampling technique for evaluation of the uncertainty in fragility estimation. Narrower confidence intervals were obtained for fragility curves versus Psa than versus PGA due to the higher correlation of the former with structural response.

We note that the KDE approach can also be applied for estimation of seismic fragilities in terms of two or more intensity measures of the seismic excitation. In the recent years, *fragility surfaces* have emerged as a promising tool for representing seismic vulnerability as a function of two intensity measures. The present study opens new paths for estimation of fragility surfaces as well.

REFERENCES

[1] K. Mackie and B. Stojadinovic. Improving probabilistic seismic demand models through refined intensity measures. In *Proc. 13th World Conf. Earthquake Eng.* Int. Assoc. for Earthquake Eng. Japan, 2004.

[2] B. Sudret and C.-V. Mai. Computing seismic fragility curves using polynomial chaos expansions. In G. Deodatis, editor, *Proc. 11th Int. Conf. Struct. Safety and Reliability (ICOSSAR'2013), New York, USA*, 2013.

[3] B. Sudret and C.-V. Mai. Calcul des courbes de fragilité par approches non-paramétriques. In *Proc. 21^e Congrès Français de Mécanique (CFM21), Bordeaux*, 2013.

[4] Pacific Earthquake Engineering and Research Center. *OpenSees: The Open System for Earthquake Engineering Simulation*, 2004.

[5] K. Mackie and B. Stojadinovic. Seismic demands for performance-based design of bridges. Technical report, Pacific Earthquake Engineering Research Center, 2003.

[6] T.-H. Kim, K.-M. Lee, Y.-S. Chung, and H.M. Shin. Seismic damage assessment of reinforced concrete bridge columns. *Engineering Structures*, 27(4):576–592, 2005.

[7] M. Shinozuka, M. Feng, H. Kim, and S. Kim. Nonlinear static procedure for fragility curve development. *J. Eng. Mech.*, 126(12):1287–1295, 2000.

(a) Maximum likelihood estimation approach

(b) Kernel density estimation approach

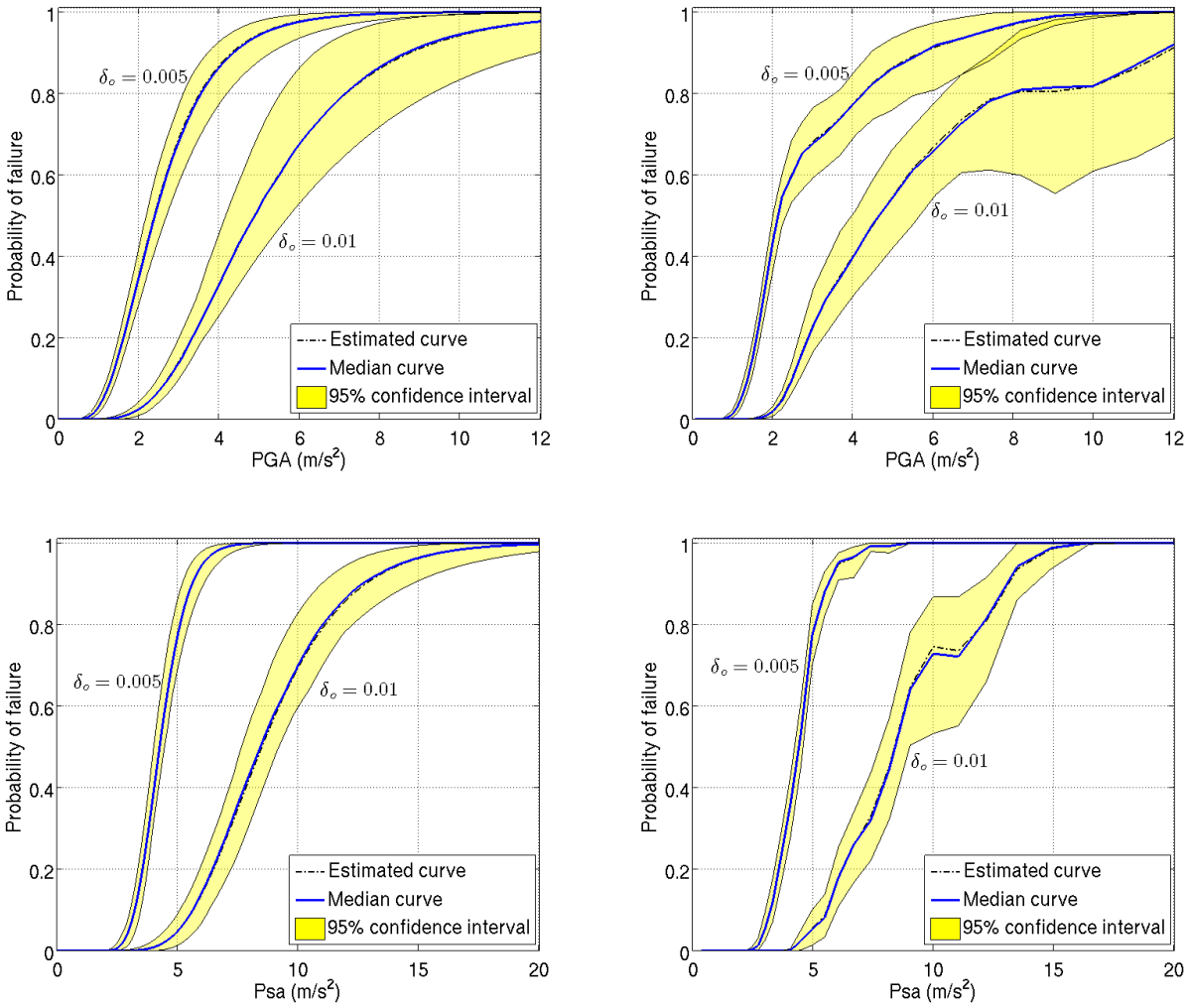

Figure 10: Estimated, median bootstrap fragility curves and 95% confidence intervals by maximum likelihood estimation and kernel density estimation approaches. The median curves and 95% confidence intervals are computed using bootstrap re-sampling technique with 100 replications.

[8] D. C. Kent and R. Park. Flexural members with confined concrete. *J. Struct. Div.*, 97(7):1969–1990, 1971.

[9] J. B. Mander, M. J. Priestley, and R. Park. Theoretical stress-strain model for confined concrete. *J. Struct. Eng.*, 114(8):1804–1826, 1988.

[10] M. L. Marsh and S. J. Stringer. *Performance-based Seismic Bridge Design, A Synthesis of Highway Practice*, volume 440. Transportation Research Board, Washington D.C., 2013.

[11] Y. Hose, P. Silva, and F. Seible. Development of a performance evaluation database for concrete bridge components and systems under simulated seismic loads. *Earthquake Spectra*, 16(2):413–442, 2000.

[12] *Visual Catalog of Reinforced Concrete Bridge Damage*. Office of Structures Maintenance and Investigation, California Department of Transportation, 2006.

[13] M. P. Berry and M. O. Eberhard. *Performance modeling strategies for modern reinforced concrete bridge columns*. Number 07. Pacific Earthquake Engineering Research Center, 2007.

[14] Y. Lu, X. Gu, and J. Guan. Probabilistic drift limits and performance evaluation of reinforced concrete columns. *J. Struct. Eng.*, 131(6):966–978, 2005.

[15] T. Rossetto and A. Elnashai. Derivation of vulnerability functions for european-type RC structures based on observational data. *Engineering Structures*, 25(10):1241 – 1263, 2003.

[16] L. Eads, E. Miranda, H. Krawinkler, and D. Lignos. An efficient method for estimating the collapse risk of structures in seismic regions. *Earthquake Eng. Struct. Dyn.*, 42(1):25–41, 2013.

[17] M. Shinozuka, M. Feng, J. Lee, and T. Naganuma. Statistical analysis of fragility curves. *J. Eng. Mech.*, 126(12):1224–1231, 2000.

[18] M. Wand and M.C. Jones. *Kernel smoothing*. Chapman and Hall, 1995.

[19] B. Sudret, C.-V. Mai, and K. Konakli. Computing seismic fragility curves using non-parametric representations. *Earthquake Eng. Struct. Dyn. (submitted for publication)*, 2014.

[20] J. Padgett, B. Nielson, and R. DesRoches. Selection of optimal intensity measures in probabilistic seismic demand models of highway bridge portfolios. *Earthquake Eng. Struct. Dyn.*, 37(5):711–725, 2008.

[21] S. Jankovic and B. Stojadinovic. Probabilistic performance based seismic demand model for R/C frame buildings. In *Proc. 13th World Conf. Earthquake Eng.*, 2004.

[22] B. Efron. *The jackknife, the bootstrap and other resampling plans*, volume 38. SIAM, 1982.