Construction of seismic fragility curves using the Karhunen-Loève expansion

Fabien Giraudeau
LaMSID (UMR EDF-CNRS-CEA 8193), France

Michel Fogli
Université Blaise Pascal de Clermont-Ferrand, France

Irmela Zentner
EDF R&D, France

Cyril Feau
CEA, DEN, DM2S, SEMT, EMSI, F-91191 Gif-sur-Yvette, France

SUMMARY:
When performing safety studies, the seismic vulnerability of industrial plants is often characterized by a *fragility curve*. Such a curve expresses the failure probability of a structure as a function of seismic intensity. The fragility curve is often supposed to follow a lognormal distribution. Furthermore, the models used to describe the ground motion do not properly reproduce the natural variability, and a scaling method is often used to account for different seismic levels. In this study, we propose to avoid these simplifications and to study their impact on the fragility curves. In particular, we use a new model for seismic ground motion based on a Karhunen-Loève expansion that is easily adjustable to real data. Then, a Karhunen-Loève expansion is used in order to reduce the number of mechanical analysis for determining fragility curves.

Keywords: Seismic signals, fragility curves, Karhunen-Loève, clustering.

1. INTRODUCTION

When performing safety studies, the seismic vulnerability of industrial plants is often characterized by a *fragility curve*. Such a curve expresses the failure probability of a structure as a function of seismic intensity.

The numerical construction of a fragility curve involves a large number of mechanical analyses, performed for different seismic intensities. However, the number of recorded accelerograms available in databases is generally not sufficient to conduct such studies and we have to simulate artificial ones.

In this article, a new method recently proposed by Zentner and Poirion (Zentner *et al.*, 2012) is used for generating artificial non stationary ground motion time histories. This method is based on the Karhunen-Loève expansion and produces artificial accelerograms whose statistical properties are close to the real ones.

The aim of our work is to develop numerically efficient methods for evaluating fragility curves without having to scale the accelerograms or to consider lognormal fragility curves, assumption classically used in parametric studies.

The scaling method supposes that the frequency content of the excitation does not depend on its intensity. It consists in multiplying the reference accelerograms by a coefficient to evaluate the failure probabilities corresponding to the *scaled intensities*.

The Karhunen-Loève expansion allows to simulate samples of accelerograms for a given site specific scenario. These artificial accelerograms are partitioned in classes of growing intensity and the failure probability is evaluated for each class. In order, to reduce the number of mechanical analysis, we propose to use the Karhunen-Loève (KL) expansion also for the structural responses. Thus, for each class, a non stationary and non Gaussian process is identified from a “small” number of structural responses. Then a sample of structural responses is simulated using the KL expansion. This allows estimating failure probabilities and constructing fragility curves without scaling or lognormal hypothesis.

2. MODELISATION OF THE SEISMIC EXCITATION

The ground acceleration resulting from an earthquake is modelled as a R-valued second-order zero-mean stochastic process $\Gamma = \{\Gamma(t)\}_{t\in[0,T]}$ based on a probability space ($\mathcal{A},\mathcal{F}$,P) and indexed on the interval $[0,T]$, $T > 0$, where T is the duration of the earthquake. Furthermore, this process is assumed to be mean-square continuous, with continuous trajectories (Soize, 1993). A seismic event (i.e. a ground acceleration record) $\{\gamma(t)\}_{t\in[0,T]}$ can then be considered as a particular realization of this process. In practice, such event can be characterized by several indicators:

- the PGA (Peak Ground Acceleration), defined by $PGA = \max_{t\in[0,T]}|\gamma(t)|$;
- the PSA (Pseudo Spectral Acceleration), given by $PSA(f_0,\chi_0) = (2\pi f_0)^2 \max_{t\in[0,T]}|y(t;f_0,\chi_0)|$, where $\{y(t;f_0,\chi_0)\}_{t\in[0,T]}$ is the response of a single-degree-of-freedom (SDOF) linear oscillator, with natural frequency f_0 and damping ratio χ_0, submitted to the external excitation $\{\gamma(t)\}_{t\in[0,T]}$. This indicator integrates the effect of the structure and verifies: $\lim_{f_0\to+\infty} PSA(f_0,\chi_0) = PGA$, also called ZPA (Zero Period Acceleration);
- the RS (Response Spectrum) at fixed damping ratio χ_0, defined as the function $RS : f_0 \to RS(f_0) = PSA(f_0,\chi_0)$.

Numerous seismic models can be found in the literature (see, for example, Kanai, 1957, Tajimi, 1960, Clough *et al.*, 1975, Rezaian *et al.*, 2010). In this work, we use the probabilistic model proposed by Zentner and Poirion (Zentner *et al.*, 2012). Based on a Karhunen-Loève expansion of the seismic excitation, this non Gaussian and non stationary model is easily identifiable from real data and can easily be simulated.

In what follows, H denotes the Hilbert space $L^2_R([0,T],dt)$ of the (classes of) mappings from $[0,T]$ into R, square integrable with respect to the Lebesgue measure dt, this space being endowed with the usual inner product $\langle f\,|\,g\rangle_H = \int_0^T f(t)g(t)dt$.

Karhunen-Loève expansion

Let $X = \{X(t)\}_{t\in[0,T]}$ be a R-valued second-order zero-mean stochastic process based on (*A*,*F* ,P), indexed on $[0,T]$, mean-square continuous, with continuous trajectories, and let $R_X : [0,T]^2 \to R : (t,t') \to R_X(t,t') = E[X(t)X(t')]$ and $Q_X : H \to H : f \to Q_X f$ be respectively its continuous autocorrelation function and its autocorrelation kernel such that, $\forall (f,t) \in H\times[0,T]$, $(Q_X f)(t) = \int_0^T R_X(t,t')f(t')dt'$. In the above expression, $E[\cdot]$ designs the mathematical expectation.

The autocorrelation kernel Q_X is a Hilbert-Schmidt positive self-adjoint continuous linear operator. As a result, according to the Mercer theorem (Mercer, 1909):

1. the spectrum of its eigenvalues $(\lambda_\alpha)_{\alpha\geq 1}$ is countable and form a positive monotonic decreasing sequence converging to 0;
2. the associated eigenfunctions $\{\phi_\alpha\}_{\alpha\geq 1}$, such that, $\forall t \in [0,T]$:

$$(Q_X\phi_\alpha)(t) = \int_0^T R_X(t,t')\phi_\alpha(t')dt' = \lambda_\alpha\phi_\alpha(t) \qquad (2.1)$$

are continuous on $[0,T]$ and form a Hermitian basis of H (*i.e.* they are orthonormal for the inner product $\langle .|. \rangle_H$ and the vector space spanned by these functions is dense in H);

3. R_X has the following representation:

$$R_X(t,t') = \sum_{\alpha=1}^{+\infty} \lambda_\alpha\phi_\alpha(t)\phi_\alpha(t') \qquad (2.2)$$

where the convergence in Eq.(2.2) is absolute and uniform.

From this result and the Karhunen-Loève theorem (Karhunen, 1947, Loève, 1978), there exists then a countable family $\{\xi_\alpha\}_{\alpha\geq 1}$ of uncorrelated R-valued second-order random variables based on (*A*,*F* ,P), with zero mean and unit variance, given by:

$$\xi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}}\langle X|\phi_\alpha\rangle_H = \frac{1}{\sqrt{\lambda_\alpha}}\int_0^T X(t)\phi_\alpha(t)dt \qquad (2.3)$$

such that the sequence $\left(\sum_{\alpha=1}^{M}\sqrt{\lambda_\alpha}\xi_\alpha\phi_\alpha(t)\right)_{M\geq 1}$ converges in mean-square to $X(t)$, uniformly in t. As a result, the process X has the following representation, referred to as the Karhunen-Loève expansion:

$$X(t) = \sum_{\alpha=1}^{+\infty} \lambda_\alpha\xi_\alpha\phi_\alpha(t) \qquad (2.4)$$

Based on this result, a similar representation has been chosen for the seismic excitation process $\Gamma = \{\Gamma(t)\}_{t\in[0,T]}$, but with only a finite number M of terms in the representation (2.4) (corresponding to the M first eigenvalues and eigenfunctions $(\lambda_\alpha,\phi_\alpha)_{1\leq\alpha\leq M}$ of the autocorrelation function R_Γ of Γ) and assuming that the random variables $(\xi_\alpha)_{1\leq\alpha\leq M}$ are mutually independent. In other words, we have chosen for this process the following model:

$$\Gamma(t) = \sum_{\alpha=1}^{M} \lambda_\alpha\xi_\alpha\phi_\alpha(t) \qquad (2.5)$$

where $(\xi_\alpha)_{1\leq\alpha\leq M}$ is a M-family of mutually independent R-valued second-order random variables based on (*A*,*F* ,P), with zero mean and unit variance, given by:

$$\xi_\alpha = \frac{1}{\sqrt{\lambda_\alpha}}\int_0^T \Gamma(t)\phi_\alpha(t)dt \qquad (2.6)$$

and $(\lambda_\alpha, \phi_\alpha)_{1\le\alpha\le M}$ is the family of the M first solutions (ranked by decreasing values of λ) of the spectral problem: find $(\lambda,\phi)\in R_+ \times H$ such that, $\forall t \in [0,T]$,

$$\int_0^T R_\Gamma(t,t')\phi(t')dt' = \lambda\phi(t) \qquad (2.7)$$

This model was identified from the available seismic signals using the following procedure.
Let $\{\gamma^{(l)},\ 1\le l\le L\}$, $L\ge 1$, be the L-family of these signals, in which each element $\gamma^{(l)} = \{\gamma^{(l)}(t)\}_{t\in[0,T]}$ is regarded as a trajectory of Γ. In practice, each signal $\gamma^{(l)}$ is given in the form of a finite family of real numbers $\{\gamma^{(l)}(t_j)\}_{1\le j\le N}$ corresponding to the recorded values of $\gamma^{(l)}$ at points $\{t_j\}_{1\le j\le N}$ of a regular partition $0 = t_1 < t_2 < ... < t_N = T$ of $[0,T]$. By assimilating these L sampled experimental signals $\left(\{\gamma^{(l)}(t_j)\}_{1\le j\le N}\right)_{1\le l\le L}$ to L numerical realizations of Γ and using an appropriate statistical estimator it is then possible to obtain an estimate of the autocorrelation function R_Γ. Let $\{\tilde{R}_\Gamma(t_i,t_j)\}_{1\le i,j\le N}$ be this estimate, such that, $\forall(i,j)\in\{1,...,N\}^2$:

$$\tilde{R}_\Gamma(t_i,t_j) = \frac{1}{L}\sum_{l=1}^{L}\gamma^{(l)}(t_i)\gamma^{(l)}(t_j) \qquad (2.8)$$

Substituting it in Eq.(2.7) and solving this equation by a Galerkin method gives the M sought solutions $\left(\tilde{\lambda}_\alpha, \{\tilde{\phi}_\alpha(t_j)\}_{1\le j\le N}\right)_{1\le\alpha\le M}$. Inserting these solutions and the successive experimental trajectories of Γ into Eq.(2.6) leads to $L\times M$ integrals whose numerical calculation using the partition of $[0,T]$ considered above gives:

$$\tilde{\xi}_\alpha^{(l)} = \frac{1}{\sqrt{\tilde{\lambda}_\alpha}}\sum_{j=1}^{N}\gamma^{(l)}(t_j)\tilde{\phi}_\alpha(t_j)\Delta t,\ 1\le l\le L, 1\le\alpha\le M \qquad (2.9)$$

where Δt is the time step. We obtain in this way, for each random variable ξ_α, a L-sample of approximated realizations $\{\tilde{\xi}_\alpha^{(l)}\}_{1\le l\le L}$, from which it is then possible to statistically estimate the cumulative distribution function (CDF) or the probability density function (PDF) of ξ_α. At this stage, each random variable ξ_α is therefore specified by an estimate of its distribution (CDF or PDF). From these estimated distributions we can simulate the random variables $(\xi_\alpha)_{1\le\alpha\le M}$, then, using Eq.(2.5), the process Γ itself. It should be noticed that the proposed model (2.5) is non Gaussian and non stationary. Furthermore, it is easy to simulate, readily identifiable from seismic records and can be continuously enriched.

3. CONSTRUCTION OF THE FRAGILITY CURVES

The probability of failure of a structure subject to a seismic excitation $\Gamma = \{\Gamma(t)\}_{t\in[0,T]}$ can be expressed as the probability of threshold crossing of a level b, by one of the response parameters of the structure, denoted by $Y = \{Y(t)\}_{t \in [0,T]}$. In seismic engineering, the scaling method is often used for evaluating the fragility curves. This method involves three steps:

- simulate N_s signals with PGA (respectively PSA) denoted $(PGA_i)_{i \in \{1,\ldots,N_s\}}$ (respectively $(PSA_i)_{i \in \{1,\ldots,N_s\}}$) from the same model;
- normalise the signals by $\frac{1}{N_s}\sum_{i=1}^{N_s} PGA_i$ (respectively $\frac{1}{N_s}\sum_{i=1}^{N_s} PSA_i$) and then multiply them by the scaling coefficient;
- compute the conditional probability of failure $P_{ZPA}\left(\max_{t \in [0,T]} |Y(t)| > b\right)$ (respectively $P_{PSA}\left(\max_{t \in [0,T]} |Y(t)| > b\right)$).

As already mentioned, this method supposes that the frequency content of the excitation does not depend on its intensity. In this work, we want to avoid this assumption and assess its influence on the fragility curve. For the model proposed in section 2, it is necessary to gather the signals in classes, what can be done according to different methods (see section 4). For each class, the curve represents the conditional probability $P_{PGA}\left(\max_{t \in [0,T]} |Y(t)| > b\right)$ (respectively $P_{PSA}\left(\max_{t \in [0,T]} |Y(t)| > b\right)$) as a function of the mean PGA (respectively PSA) of the class. The fragility curves determined in the following section, using the KL expansion, will be approximated (in the least-squares sense) by lognormal curves in order to evaluate the relevance of this common assumption. Moreover, aiming at industrial applications, it is necessary to reduce the number of mechanical analysis. But, if this number is decreased, the size of each class is also reduced and the proposed method loses accuracy. For classes where only a few structural responses are available, it will be necessary to enrich them by applying the KL expansion to the initial population in order to increase the sample size and to evaluate failure probabilities.

4. APPLICATIONS

Reference signals and clustering

The database of real accelerograms used to generate artificial time histories is a subset of 97 signals of the European Strong-Motion Database (ESD, 2000). This subset is chosen according to magnitudes 5.5 < M < 6.5 and distances in the range 0 < D < 20 km. Using the KL method, 30000 new signals have been simulated and clustered by PGA and by PSA for $f_0 = 8\,\text{Hz}$ and $\chi_0 = 5\%$. The clustering was done using the k-means method (Jain, 1999). We thus obtained 30 classes, each containing a minimum of one hundred signals. Figure 1 shows the evolution of the mean RS (response spectrum) with respect to f_0 for each class of PGA and PSA.

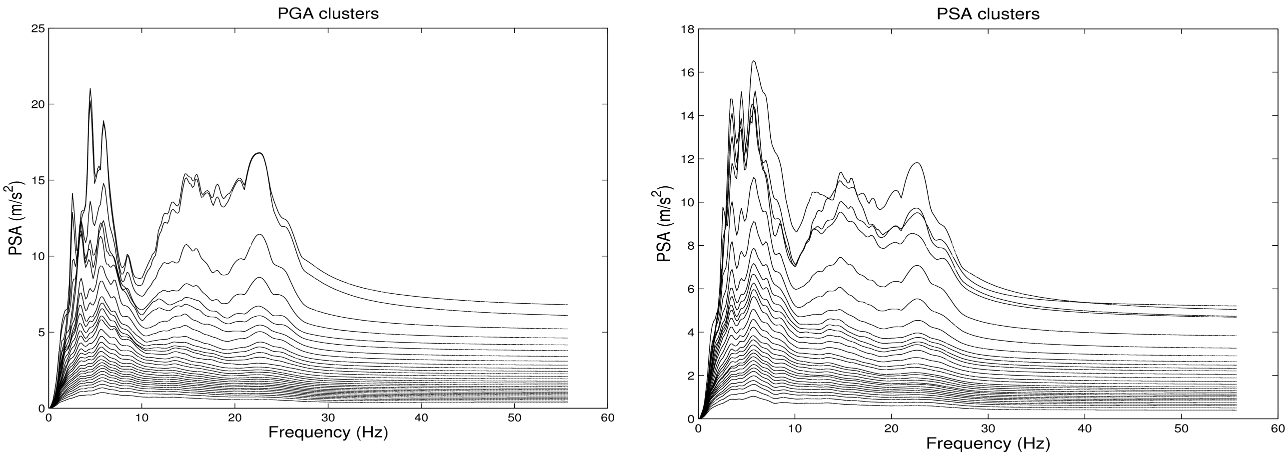

Figure 1. Evolution of the mean RS according to f_0 for each class of PGA and PSA signals simulated for 30 classes of PGA, and 30 classes of PSA for 8 Hz and 5%.

The frequency contents differ between the PGA (or PSA) classes. This result shows the simplification introduced by the scaling method, which supposes the same frequency contents for each of the excitation levels.

Mechanical models

For the numerical applications two simple mechanical models are chosen: a SDOF linear oscillator and a SDOF nonlinear oscillator. These models represent global behaviours of simple structures whose responses are mainly on the first eigenmode. In both cases, we choose a natural frequency $f_0 = \omega_0 / 2\pi = 8\text{Hz}$ and a damping ratio $\chi_0 = 5\%$.

The dynamics of the linear and nonlinear oscillator are respectively governed by the following equations:

$$\ddot{y}(t) + 2\chi_0\omega_0\dot{y}(t) + \omega_0^{\;2} y(t) = -\gamma(t) \qquad (4.1)$$

$$\ddot{y}(t) + 2\chi_0\omega_0\dot{y}(t) + F(y(t),\lambda) = -\gamma(t) \qquad (4.2)$$

where $\{\gamma(t)\}_{t\in[0,T]}$ is the acceleration and $\omega_0 = \sqrt{k/m}$, with $m = 1\,\text{kg}$. The non linearity $F(y(t),\lambda)$ is defined on Figure 2. It describes a kinematic hardening phenomenon.

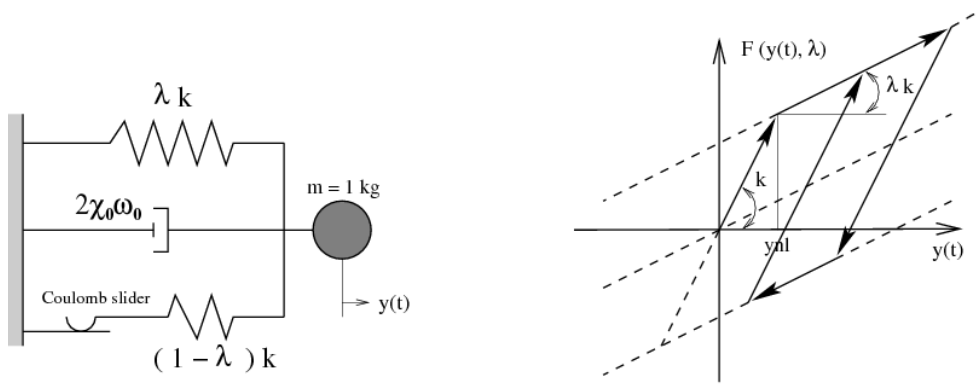

Figure 2. Rheological model of the oscillator with kinematic hardening.

For the calculations, we have chosen the following value for the slope ratio: $\lambda = 0.2$. The nonlinear oscillator is designed for having a linear behaviour until the 10^{th} class (in which $RS(f_0) = 2.39\ m/s^2$), when the signals are clustered by PGA. Its elastic limit is $y_{nl} = 2.39/(2\pi f_0)^2 = 9.56\times10^{-4}$ m. For a PSA clustering, this corresponds to the 11^{th} class. The failure threshold b is chosen such that the displacement ductility demand $\mu = b/y_{nl}$ equals 2. Thus, the threshold b equals 1.9×10^{-3} m. We have arbitrarily decided to keep the same threshold for the linear oscillator. Besides, for both models, the initial displacement and velocity are zero.

Results obtained

The Figures 3 and 4 represent the fragility curves according to the classes of PGA and PSA for $f_0 = 8\,\text{Hz}$ and $\chi_0 = 5\%$. On these figures the results obtained with the proposed method are compared to the ones derived from the lognormal approximation and the scaling method. 95% confidence intervals are also shown. When the probability of failure p is estimated by $\hat{p} = \frac{N_{def}}{N}$ where N is the sample size and N_{def} the number of failures, then the 95% confidence interval for this estimate can be evaluated as (Saporta, 2006):

$$\left[\hat{p} - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{N}}, \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{N}}\right]$$

In order to account for uncertainty on the mechanical model, the natural frequency, the damping ratio and the slope ratio have been modelled as uniform random variables $\mathcal{U}(0.9c,1.1c)$, where c is the reference value. For the scaling method, the 1695 signals of the 10^{th} class for the PGA curve, and the 1586 signals of the 11^{th} class for the PSA curve have been multiplied by a homothetic factor.

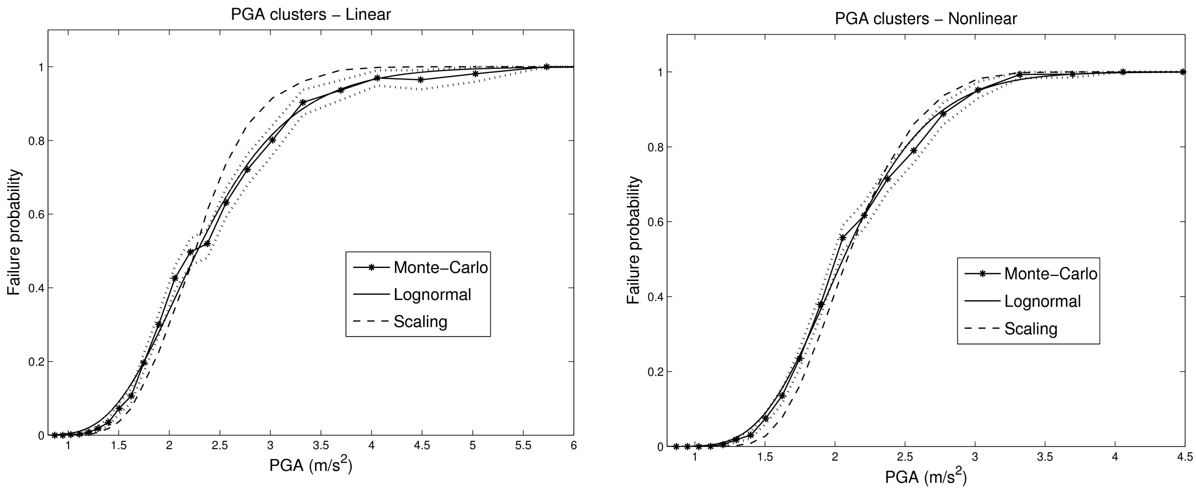

Figure 3. Fragility curves for linear and nonlinear oscillators, with 30 PGA classes. Thick points: proposed (Monte-Carlo) method with 95% confidence interval (dotted line); solid line: lognormal approximation; dashes: scaling method.

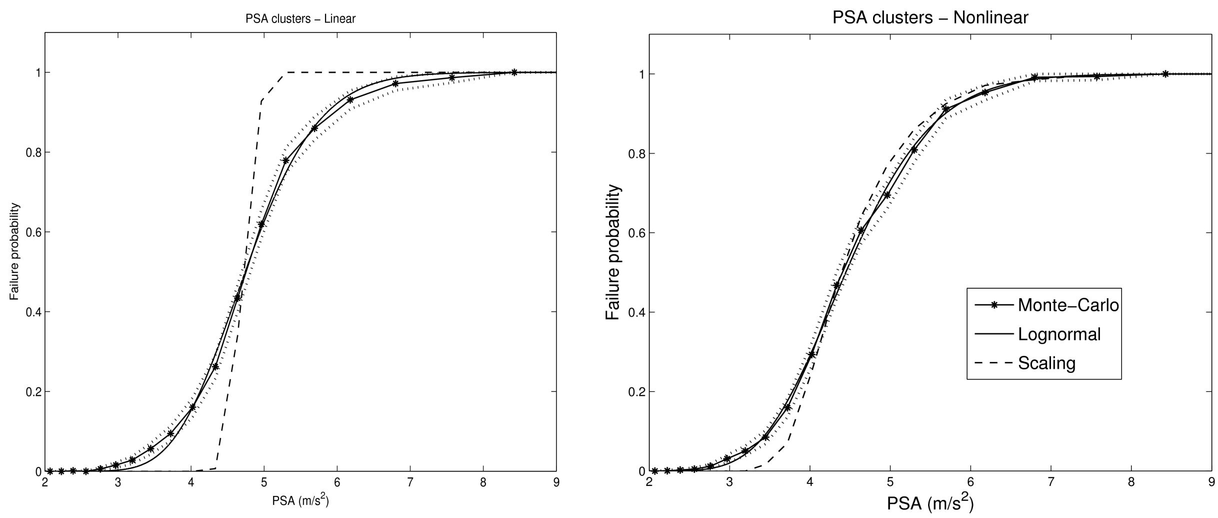

Figure 4. Fragility curves for linear and nonlinear oscillators, with 30 PSA classes. Thick points: proposed (Monte-Carlo) method with 95% confidence interval (dotted line); solid line: lognormal approximation; dashes: scaling method.

We can observe that the scaling method and the lognormal approximation lead to fragility curves which differ from those provided by the proposed method.

Karhunen-Loève expansion of the structural response

We now want to construct the fragility curves by applying the proposed method to the structural responses (i.e. by applying the KL expansion not only to the excitation but also to the seismic response of the structure). Currently, this work is in progress. We present here some preliminary results of a study dealing with the nonlinear oscillator considered above. It consists in estimating the PDF of the absolute value of the extreme structural response. This estimate was first performed using the 2051 experimental signals of the 6^{th} PGA class. Then, it was obtained from 1000 simulated signals provided by the KL expansion by taking, as basic signals for this method, firstly 50 and then 100 signals of the

6^{th} PGA class (randomly chosen in this class). In each case, the PDF was estimated using Gaussian kernel smoothing. The two estimates are presented on Figure 5.

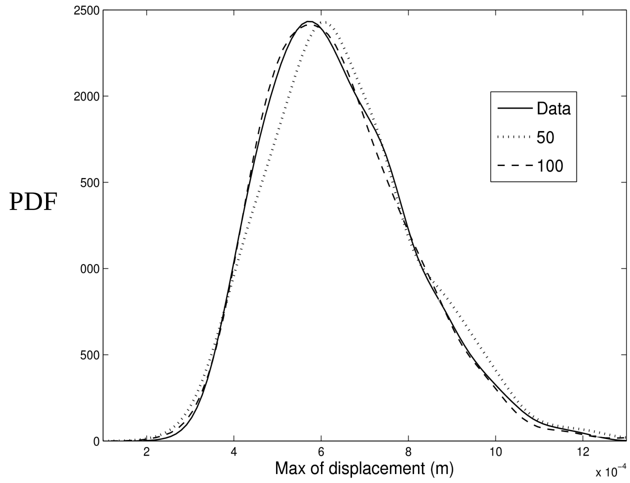

Figure 5. PDF of the absolute value of the extreme response of the nonlinear oscillator. Solid line: estimate obtained from the 2051 experimental signals of the 6^{th} PGA class; dotted and dashed lines: estimates obtained from 1000 simulated signals derived from the proposed model identified first from 50 experimental signals, (dots), then from 100 experimental signals (dashes).

We can see that the KL expansion based on 100 initial experimental signals gives a better approximation of the PDF than the same method based on 50 initial experimental signals. The precision of the estimate of the autocorrelation function of the response process, and of the random variables ξ_α, increases with the number of signals. The higher this number, the more the KL model will be representative of the 2051 initial signals of the 6^{th} PGA class. The quality of the approximation obtained encourages continuing in this direction, namely applying the KL expansion to the structural responses, in order to reduce the number of mechanical analysis.

5. CONCLUSIONS AND PERSPECTIVES

In this paper, we have proposed a methodology for constructing fragility curves where the seismic excitation is modelled by a non Gaussian and non stationary stochastic process. The model of such process is constructed by means of a ground motion database and using the Karhunen-Loève expansion. This methodology allows the determination of fragility curves without scaling and without assuming an *a priori* lognormal fragility curve. In order, to reduce the number of mechanical analysis, it has been proposed to use the Karhunen-Loève expansion for the structural responses. Currently, this work is in progress nevertheless, the first results obtained encourage continuing in this direction.

6. REFERENCES

Clough R., Penzien J. (1975), Dynamics of structures, McGraw-Hill.

ESD (European Strong-Motion Database) (2000). http://www.isesd.hi.is/ESD_Local/frameset.htm

Jain A., Murty M., Flynn P. (1999). Data clustering: a review. Journal ACM Computing Surveys.

Kanai K. (1957). Semi-empirical formula for the seismic characteristics of the ground motion. Bulletin of the Earthquake Research Institute, University of Tokyo, **Vol. XXXV-2:** p. 308-325.

Karhunen K. (1947). Über lineare Methoden in der Wahrscheinlichkeitsrechnung. Ann. Acad. Sci. Fennicae. Ser.A.I. Math.-Phys. **Vol. XXXVII:** p. 1-79.

Loève M. (1978). Probability theory. Graduate Texts in Mathematics, 4th edt. Springer –Verlag. Vol. II. ISBN 0-387-90262-7.

Mercer J. (1909). Functions of positive and negative type and their connection with the theory of integral equations. Phils. Trans. Roy. Soc. London. **Vol. A CCIX.**

Rezaian S., Der Kiureghian A.(2010). Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering and Structural Dynamic*. **Vol. XXXIX**: p. 1155-1180.

Saporta G. (2006). Probabilités, Analyse de Données et Statistiques – 2nde édition. Editions TECHNIP, Paris.

Soize C. (1993), Méthodes mathématiques en analyse du signal, Masson.

Tajimi H. (1960). A statistical method of determining the maximum response of a building structure during an earthquake. 2nde WCEE, Tokyo. p. 781-798.

Zentner I., Poirion F. (2012). Enrichment of seismic ground motion databasis using Karhunen-Loève expansion. *Earthquake Engineering and Structural Dynamics*. 10.1002/eqe.2166.