16th World Conference on Earthquake, 16WCEE 2017
Santiago Chile, January 9th to 13th 2017
Paper N° 595
Registration Code: S-B1462715524

STRUCTURAL RELIABILITY ANALYSIS USING DETERMINISTIC HYBRID SIMULATIONS AND ADAPTIVE KRIGING METAMODELING

G. Abbiati(1), R. Schöbi(2), B. Sudret(3), B. Stojadinovic(4)

(1) *Postdoctoral Researcher, Department of Civil, Environmental and Geomatic Engineering (D-BAUG), IBK, ETH Zurich, Switzerland, abbiati@ibk.baug.ethz.ch*
(2) *Ph.D. Candidate, Department of Civil, Environmental and Geomatic Engineering (D-BAUG), IBK, ETH Zurich, Switzerland, schoebi@ibk.baug.ethz.ch*
(3) *Professor, Department of Civil, Environmental and Geomatic Engineering (D-BAUG), IBK, ETH Zurich, Switzerland, sudret@ibk.baug.ethz.ch*
(4) *Professor, Department of Civil, Environmental and Geomatic Engineering (D-BAUG), IBK, ETH Zurich, Switzerland, stojadinovic@ibk.baug.ethz.ch*

Abstract
Reliability analysis aims at determining the probability of failure of a stochastic system and typically relies on purely numerical simulations. Nevertheless, trusting mathematical models up to failure can be questioned, especially for structural components characterized by a strongly nonlinear response. From this perspective, hybrid simulation is as a suitable candidate to replace pure numerical time-history analyses in the process of structural reliability assessment. Arguably, hybrid simulation suffers the same limitations as the large finite element models, making non-intrusive reliability methods based on metamodeling the most attractive solution strategy for hybrid simulation based reliability analysis. Along these lines, this paper explores a combination of an active learning reliability method based on Kriging metamodeling and hybrid simulation, which is used to provide the real structural response. A numerical validation of the proposed approach is presented for a nonlinear two-degrees-of-freedom series system. The selected benchmark case study covers a wide class of structures whose response can be evaluated using hybrid simulation.

Keywords: Hybrid Simulation, Kriging Metamodeling, Gaussian Process, Active Learning, Adaptive Experimental Design.

1 Introduction
Reliability Analysis (RA) aims to determine the probability of failure of a stochastic system. A performance function defines the system state, i.e. failure or not failure, in the overall probability space of the random input parameters. The probability integral over the failure domain corresponds to the probability of failure [1]. In the current practice, the probability of failure is estimated using pure numerical simulations, assuming that the behavior of the structure is well-known and correctly modeled up to failure, at least in a probabilistic sense. This assumption can be questioned, especially for structures with components characterized by a strongly nonlinear response and a lack of a reliable numerical model, namely in the presence of epistemic uncertainty [2]. From this perspective, hybrid simulation (HS) is a suitable candidate to replace pure numerical time-history response analyses in the process of structural RA.

A hybrid model of the prototype structural system combines numerical and physical substructures (NSs and PSs). The PS of the hybrid model is tested in the laboratory precisely because of a lack of reliable mathematical models, while the NS is instantiated in structural analysis software. The dynamic response of the hybrid model is predicted using a time-stepping response history analysis. In a typical hybrid simulation, a computer-controlled system applies displacements to the PS using servo-hydraulic actuators and feeds back the corresponding restoring forces to the time integration algorithm where the next solution step is computed. When the response of the PS does not depend on the rate of loading, hybrid simulation can be performed at an extended time scale, typically 50-200 times slower than the actual earthquake rate, requiring inertia and damping forces to be modeled numerically. This is the so-called Pseudodynamic (PsD) testing method that allows for improving the quality of the test by increasing the signal-to-noise ratio of response signals and reducing the control tracking error [3,4,5].

Arguably, HS suffers the same limitations as those of large Finite Element (FE) models [6], as well as additional limitations associated with accuracy and duration of experimental testing. Therefore, crude Monte Carlo Simulation (MCS) is not affordable, making non-intrusive solution algorithms based on metamodeling the most appealing strategy for HS-based RA. From this perspective, this paper explores a combination of the active learning reliability method based on Kriging metamodeling [7,8] and HS, which is used to sample the limit state function and provide real structural response data near and at the limit state. An important objective is to minimize the size of the training set because each element involves physically tested, and possibly damaged, PSs. To this end, Sobol low-discrepancy sequences are selected to provide the initial pool of random input parameter samples, which fill the entire probability space [9]. Subsequent samples are selected after each HS to minimize the uncertainty of the limit state surface localization. After a predetermined number of iterations, the Kriging model is considered accurate enough to calculate probability of failure or quantile estimates using MCS.

The paper outline follows. First, the Kriging metamodeling theory is presented. The active learning reliability method is illustrated afterwards. On this basis, a reliability method is derived that considers HS as a non-intrusive way of sampling the limit state function of the system being investigated. The resulting Adaptive Kriging HS reliability method, AK-HS hereinafter, is validated using a purely numerical analysis of a nonlinear two-degrees-of-freedom (2-DOF) series system. The selected benchmark case study covers a wide class of structures whose response can be evaluated using hybrid simulation: the dynamic response of these structures is governed by a few response modes and a few localized nonlinearities. The stiffness parameters of the boundary springs in the 2-DOF series system, which belong to the NS, are selected as input random variables. A limit state function is defined with respect to the maximum absolute displacement achieved by the PS hysteretic spring. The sensitivity of the computed failure probability magnitude to the nonlinearity in the dynamic response is investigated considering a reasonable number of deterministic HS. Results, characterized by fast convergence of statistical estimates, are presented and discussed.

2 Failure probability and quantile estimation

Given a probabilistic model corresponding to an m-dimensional random vector $\boldsymbol{X}$ with probability density function $f_{\boldsymbol{X}}$ and a computational model $\mathcal{M}$, failure is defined as the event $F = \{\mathcal{M}(\boldsymbol{X}) \geq y_{adm}\}$ and the failure probability reads:

$$p_f = \mathbb{P}\left(\{\mathcal{M}(\boldsymbol{X}) \geq y_{adm}\}\right) = \int_{\mathcal{D}_f = \{\boldsymbol{x} \in \mathbb{R}^m : \mathcal{M}(\boldsymbol{x}) \geq y_{adm}\}} f_{\boldsymbol{X}}(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} \tag{1}$$

In the context of RA, the computational model $\mathcal{M}$ represents the Limit State Function (LSF). Due to the generally complex shape of the failure domain $\mathcal{D}_f$, the integration operation cannot be solved analytically. However, a numerical estimate of the failure probability can be obtained by Monte Carlo simulation (MCS). Given a large set of samples of the input vector $\boldsymbol{X}$, denoted by $\mathcal{S} = \{\boldsymbol{x}_1, \dots \boldsymbol{x}_n\}$, the failure probability reads:

$$\hat{p}_f = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\mathcal{M}(x) \geq y_{adm}}(\boldsymbol{x}_i) \tag{2}$$

where $\mathbb{I}_{\mathcal{M}(x) \geq y_{adm}}$ is the indicator failure function, which reads $\mathbb{I} = 1$ for $\boldsymbol{x}_i \in \mathcal{D}_f$ and 0 otherwise. This equation transforms the reliability analysis into a classification problem, where only the distinction of safe and failure domain is of interest.

A related problem is the estimation of quantiles q_α, which are defined as,

$$\mathbb{P}(\mathcal{M}(\boldsymbol{X}) \geq q_\alpha) = 1 - \alpha \tag{3}$$

with $\alpha \in]0,1[$. Assume again a large sample of the input vector $\mathcal{S}$ and the corresponding response values y_i. When ranking the response values in descending order such that $y_{(1)} \geq y_{(2)} \geq \cdots \geq y_{(n)}$, a quantile can be estimated by:

$$\hat{q}_\alpha = y_{\lfloor n(1-\alpha)\rfloor}, \tag{4}$$

Where $\lfloor n(1-\alpha)\rfloor$ is the largest integer smaller than $n(1-\alpha)$. Again, this is a classification problem, where the failure domain is defined as $\mathcal{D}_f = \{\boldsymbol{x}: \mathcal{M}(\boldsymbol{x}) \geq q_\alpha\}$.

However, the main drawback of these classification methods is that they rely on a large sample set $\mathcal{S}$ to estimate accurately the quantity of interest. When the computational model $\mathcal{M}$ is expensive to evaluate, such as in HS, the analysis becomes intractable. In order to make these analyses tractable, metamodels are introduced in the next section.

3 Reliability assessment based on hybrid simulation and Kriging metamodeling

3.1 Kriging basics

Kriging is a meta-modelling technique that considers the computational model to be a realization of a Gaussian process [10]:

$$\widehat{\mathcal{M}}(\boldsymbol{x}) = \boldsymbol{\beta}^T \boldsymbol{f}(\boldsymbol{x}) + \sigma^2 Z(\boldsymbol{x}, \omega), \tag{5}$$

where $\boldsymbol{f}(\boldsymbol{x}) = [f_1(\boldsymbol{x}), \dots, f_p(\boldsymbol{x})]$ are regression functions, $\boldsymbol{\beta}$ is a vector of coefficients, which compose the mean value of a Gaussian process. σ^2 is the corresponding variance. $Z(\boldsymbol{x}, \omega)$ is a zero-mean, unit-variance, stationary Gaussian process, which is characterized by an autocorrelation function $R(|\boldsymbol{x} - \boldsymbol{x}'|; \boldsymbol{\rho})$ and its hyper-parameters $\boldsymbol{\rho}$. The Kriging model is trained with a set of realizations $\boldsymbol{\mathcal{X}} = \{\boldsymbol{\chi}^{(i)}, i = 1, \dots, N\}$ and the corresponding responses of the computational model $\boldsymbol{\mathcal{Y}} = \{\mathcal{Y}^{(i)} = \mathcal{M}(\boldsymbol{\chi}^{(i)}), i = 1, \dots, N\}$, which together form the so-called Experimental Design (ED) $\{\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}}\}$. Kriging parameters are obtained by generalized least-squared solution:

$$\boldsymbol{\beta}(\boldsymbol{\rho}) = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \boldsymbol{\mathcal{Y}}, \tag{6}$$

$$\sigma_y^2(\boldsymbol{\rho}) = \frac{1}{N} (\boldsymbol{\mathcal{Y}} - \mathbf{F}\boldsymbol{\beta})^T \mathbf{R}^{-1} (\boldsymbol{\mathcal{Y}} - \mathbf{F}\boldsymbol{\beta}), \tag{7}$$

where $\mathbf{R}_{ij} = R(|\boldsymbol{\chi}^{(i)} - \boldsymbol{\chi}^{(j)}|; \boldsymbol{\rho})$ is the correlation matrix and $\mathbf{F}_{il} = f_l(\boldsymbol{\chi}^{(i)})$. In practice the correlation hyper-parameters are unknown and their values shall be inferred by e.g. maximum likelihood estimation.

Having determined the Kriging parameters, the prediction value of the computational model at a point $\boldsymbol{x} \in \mathcal{D}_X$ is a Gaussian variable with the following mean value and variance:

$$\mu_{\hat{Y}}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^T \boldsymbol{\beta} + \boldsymbol{r}(\boldsymbol{x})^T \mathbf{R}^{-1} (\boldsymbol{\mathcal{Y}} - \mathbf{F}\boldsymbol{\beta}), \tag{8}$$

$$\sigma_{\hat{Y}}(\boldsymbol{x}) = \sigma_y^2 (1 - \boldsymbol{r}(\boldsymbol{x})^T \mathbf{R}^{-1} \boldsymbol{r}(\boldsymbol{x}) + \boldsymbol{u}(\boldsymbol{x})^T (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \boldsymbol{u}(\boldsymbol{x})), \tag{9}$$

where $r_i(\boldsymbol{x}) = R(|\boldsymbol{x} - \boldsymbol{\chi}^{(i)}|; \boldsymbol{\rho})$ and $\boldsymbol{u}(\boldsymbol{x}) = \mathbf{F}^T \mathbf{R}^{-1} \boldsymbol{r}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x})$. Based on the trained Kriging model, the failure probability and the quantiles can be estimated.

3.2 Adaptive experimental design

The trained Kriging model, produced by the above stated process, is capable of exactly reproducing the points in the ED set. Indeed, the Kriging metamodel is interpolating, meaning that $\mu_{\hat{Y}}(\boldsymbol{\chi}^{(i)}) = \mathcal{M}(\boldsymbol{\chi}^{(i)}), i = 1, \dots, N$, exactly. This, however, does not mean it is an optimal model to predict the failure probability or the quantiles as RA requires. According to Eq. (2), RA is clearly a classification problem where each sample is categorized either as failed or non-failed. The purpose of the Kriging metamodel is indeed to approximate accurately the limit stat surface, which separates failure and non-failure domains. Adaptively enriching the ED set in a guided way can improve the accuracy of the quantity the trained Kriging model is used to predict. The learning rule that is used to pick samples as trade-off between exploration (of the input probabilistic space) and exploitation (RA, in this case) plays a crucial role in this game [11]. In order to extend the scope of HS to RA, we followed the developments by

Echard et al. [7] and its extension to quantiles in Schöbi et al. [8], who defined a specific learning rule for classification problems. The main steps of an Adaptive Kriging (AK) algorithm are listed here:

1) Generate a small initial ED by selecting realizations $\boldsymbol{\mathcal{X}}$ and computing the corresponding response values using the computational model $\mathcal{Y}^{(i)} = \mathcal{M}\left(\boldsymbol{\chi}^{(i)}\right)$.
2) Train a Kriging model $\widehat{\mathcal{M}}$ based on this ED.
3) Generate a large set $\mathcal{S} = \{\boldsymbol{x}_1, \dots \boldsymbol{x}_n\}$ and predict the response values of $\widehat{\mathcal{M}}$, i.e. $\mu_{\hat{Y}}(\boldsymbol{x})$ and $\sigma_{\hat{Y}}(\boldsymbol{x})$.
4) Estimate the quantity of interest. In the case of failure probabilities, the best estimate of the failure probability and a confidence interval is computed as:

$$\hat{P}_f = \mathbb{P}(\mu_{\hat{Y}}(\boldsymbol{x}) \geq y_{adm}), \tag{10}$$

$$\hat{P}_f^- = \mathbb{P}(\mu_{\hat{Y}}(\boldsymbol{x}) - 2\sigma_{\hat{Y}}(\boldsymbol{x}) \geq y_{adm}), \quad \hat{P}_f^+ = \mathbb{P}(\mu_{\hat{Y}}(\boldsymbol{x}) + 2\sigma_{\hat{Y}}(\boldsymbol{x}) \geq y_{adm}). \tag{11}$$

In the case of quantile estimation, the best estimate of the quantile and a confidence interval are computed in an analogous way (using Eq. (3)):

$$1 - \alpha = \mathbb{P}(\mu_{\hat{Y}}(\boldsymbol{x}) \geq \hat{q}_\alpha), \tag{12}$$

$$1 - \alpha = \mathbb{P}(\mu_{\hat{Y}}(\boldsymbol{x}) - 2\sigma_{\hat{Y}}(\boldsymbol{x}) \geq \hat{q}_\alpha^-), \quad 1 - \alpha = \mathbb{P}(\mu_{\hat{Y}}(\boldsymbol{x}) + 2\sigma_{\hat{Y}}(\boldsymbol{x}) \geq \hat{q}_\alpha^+). \tag{13}$$

5) Check for convergence. In the case of failure probabilities, the convergence criterion is $\left(\hat{P}_f^+ - \hat{P}_f^-\right)/\hat{P}_f \leq \epsilon_{P_f}$. In the case of quantiles, the corresponding convergence criterion is $(\hat{q}_\alpha^+ - \hat{q}_\alpha^-)/std(Y) \leq \epsilon_{q_\alpha}$. If it is not fulfilled, continue with step 6), otherwise stop here and return the last metamodel. Note that a value of $\epsilon_{P_f} = \epsilon_{q_\alpha} = 5\%$ leads to accurate results at reasonable costs [8].
6) Enrich the ED by a single sample $\boldsymbol{x}^* \in \mathcal{S}$ by maximizing the probability of misclassification. For failure probability estimation:

$$\boldsymbol{x}^* = \underset{\boldsymbol{x}_i \in \mathcal{S}}{\operatorname{argmax}}\, \Phi\left(-|\mu_{\hat{Y}}(\boldsymbol{x}) - y_{adm}|/\sigma_{\hat{Y}}(\boldsymbol{x})\right), \tag{14}$$

where $\Phi(\cdot)$ is the CDF value of a standard Gaussian variable. For the case of quantile estimation, the probability of misclassification is maximized by:

$$\boldsymbol{x}^* = \underset{\boldsymbol{x}_i \in \mathcal{S}}{\operatorname{argmax}}\, \Phi\left(-|\mu_{\hat{Y}}(\boldsymbol{x}) - \hat{q}_\alpha|/\sigma_{\hat{Y}}(\boldsymbol{x})\right). \tag{15}$$

7) Compute the corresponding response value $y^* = \mathcal{M}(\boldsymbol{x}^*)$ of the computational model and add $\{\boldsymbol{x}^*, y^*\}$ to the ED. Return to step 2).

After the termination of the iterative algorithm, estimate the failure probability or quantile with the last metamodel $\widehat{\mathcal{M}}$ via MCS. As can be argued from Eqs. (14) and (15), best candidate samples lie either close to the boundary of the failure domain or where the variance of the metamodel is high. In both cases, the probability of misclassification tends to its upper bound, which is equal to 0.5.

3.3 Description of the proposed reliability method

The reliability method proposed in this paper combines AK metamodeling and HS to evaluate the statistics that are related to the hybrid system response e.g., failure probabilities and quantiles. In detail, for a given sample $\boldsymbol{\chi}^{(i)}$ of the random input parameter vector, HS replaces a purely computational model to provide the corresponding response quantity $\mathcal{Y}^{(i)} = \mathcal{M}\left(\boldsymbol{\chi}^{(i)}\right)$. Arguably, the maximum number of laboratory experiments limits the size of the ED, which must be set beforehand. In order to cope with this constraint, the original AK procedure of Schöbi and co-workers [8], described in Subsection 3.2, is slightly adjusted. The steps of the resulting AK-HS reliability method are:

1) Run the initial pool of HSs to generate the starting ED $\{\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{Y}}\}$ with $\mathcal{Y}^{(i)} = \mathcal{M}(\mathcal{X}^{(i)})$.
2) Train a Kriging model $\widehat{\mathcal{M}}$ based on the starting ED.
3) Generate a large set $\mathcal{S} = \{\boldsymbol{x}_1, \ldots \boldsymbol{x}_n\}$ and predict the response values of $\widehat{\mathcal{M}}$, i.e. $\mu_{\hat{Y}}(\boldsymbol{x})$ and $\sigma_{\hat{Y}}(\boldsymbol{x})$.
4) Enrich the ED by picking the sample $\boldsymbol{x}^* \in \mathcal{S}$ that maximizes the probability of misclassification according to Eqs. (14-15).
5) Run the corresponding HS, which extends the ED to include the response pair $\{\boldsymbol{x}^*, y^*\}$.
6) Train a Kriging model $\widehat{\mathcal{M}}$ based on the last ED.
7) Loop between 3) and 6) for a pre-determined number of iterations.
8) Estimate the statistics of interest via MCS on the last Kriging metamodel according to Eqs. (10-13).

In order to probe the overall probability space, a Sobol sequence [9] provides the starting ED input $\boldsymbol{\mathcal{X}}$ while corresponding HSs provide the ED output $\boldsymbol{\mathcal{Y}}$. It is noteworthy that the ED follows probabilistic distributions of input parameters whereas the Kriging metamodel works in a standard normal space. Accordingly, a proper iso-probabilistic transformation must be performed on input quantities beforehand. It must be stressed that the proposed method is conceived for a peculiar class of hybrid systems whose dynamic response is governed by a few modes and a few lumped nonlinearities, which are typically confined to the PS. As will be shown in the following, these conditions systematically ensure a stable and almost un-biased estimates of both failure probabilities and quantiles with circa 30 HS. However, a significant dispersion of the corresponding confidence bounds is observed, which makes the original convergence check foreseen by Schöbi and co-workers [8] too strict to be satisfied in a reasonable number of HS experiments. This motivate our choice to establish the number of iterations, i.e. the number of HS tests, a priori, according to available resources.

4 Numerical validation of the active learning reliability method

4.1 Description of the reference case study

A numerical benchmark study is presented in the following to support the validation of the proposed RA method. A nonlinear 2-DOF series system is shown in Figure 1. The spring connecting the two masses is assumed to respond inelastically and represents the PS. The two masses of the system and the springs that connect them to the fixed supports are the NSs of the hybrid model. It is assumed that the support springs are elastic and linear. The PS spring would be tested in a laboratory, but in this study, its force-deformation response is modeled numerically using a displacement-driven Bouc-Wen model [12].

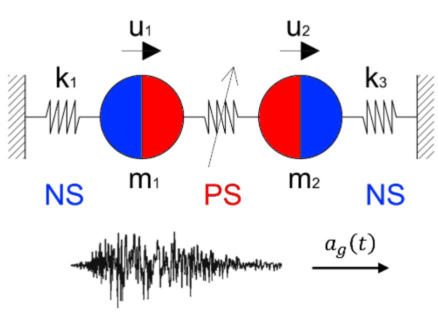


Fig. 1 - Nonlinear 2-DOF series system subjected to ground motion excitation.

The equations of motion of the hybrid model are:

$$
\begin{aligned}
\dot{u}_1 &= v_1 \\
\dot{u}_2 &= v_2 \\
\dot{v}_1 &= m_1^{-1}\big(m_1 a_g(t) + r - k_1 u_1 - c_{11} v_1 - c_{12} v_2\big) \\
\dot{v}_2 &= m_2^{-1}\big(m_2 a_g(t) - r - k_3 u_2 - c_{21} v_1 - c_{22} v_2\big) \\
\dot{r} &= \big(A - \big(\beta \operatorname{sign}\big((v_2 - v_1) r\big) + \gamma\big)|r|^n\big)(v_2 - v_1)
\end{aligned}
\tag{16}
$$

where r represents the nonlinear restoring force of the PS, u and v stand for displacement and velocity while m, c and k are mass, damping and stiffness parameters; $a_g(t)$ is the acceleration history of the selected earthquake excitation. A uniform modal damping ζ was assumed for the calculation of the damping matrix. The reference values of the parameters of the 2-DOF series hybrid model are (referring to Figure 1 and Eq. (16)):

$$m_1 = 8e3\ kg, m_2 = 9e3\ kg, k_1 = 4e5\frac{N}{m}, k_3 = 1e6\frac{N}{m},$$
$$A = 5e5\frac{N}{m}, \beta = 25, \gamma = 12.5, n = 1, \zeta = 0.04$$

The undamped eigenfrequencies of the 2-DOF hybrid model linearized about the initial (undeformed) configuration are 1.376 Hz and 2.275 Hz (i.e. the vibration mode periods are 0.727 s and 0.440 s). An accelerogram of the 1989 Loma Prieta earthquake recorded from the UCSC station [13] was selected as the reference ground motion acceleration signal. In order to observe the effects of an increasingly nonlinear response of the PS, the same reference signal was scaled to three different Peak Ground Acceleration (PGA) values. In detail, $\text{PGA}_1 = 1.00\ m/s^2$, $\text{PGA}_2 = 5.00\ m/s^2$ and $\text{PGA}_3 = 10.00\ m/s^2$ were chosen to provide a linear, a slightly nonlinear and a strongly nonlinear system response, respectively. Figure 2 depicts the un-scaled Loma Prieta accelerogram ($\text{PGA}_{ref} = 4.28\ m/s^2$) while Figure 3 compares the related response spectra after scaling to the selected PGA values.

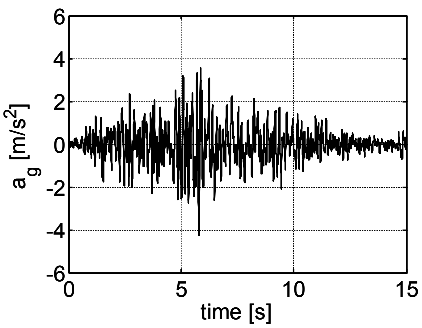


Fig. 2 - Selected Loma Prieta earthquake acceleration record ($\text{PGA}_{ref} = 4.28\ m/s^2$).

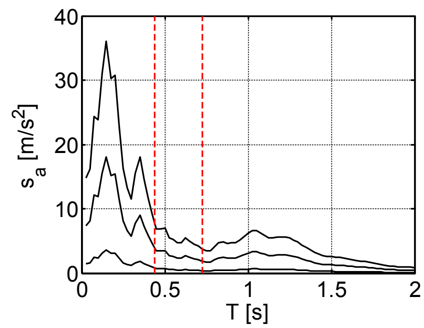


Fig. 3 - Acceleration response spectra for the selected Loma Prieta earthquake record scaled to the three PGA values.

With reference to Figure 3, dashed red lines indicate the 2-DOF hybrid model periods. Provided that $(\beta + \gamma) > 0$, the inelastic restoring force limit attained by the Bouc-Wen spring is:

$$r_y = \left(\frac{A}{\beta + \gamma}\right)^{\frac{1}{n}} \tag{17}$$

Arguably, the ratio between the maximum absolute restoring force of the Bouc-Wen spring and its inelastic force limit r_y, namely the normalized restoring force peak, is a reliable indicator of the expected degree of nonlinearity of the system response. Accordingly, it was used to check the calibration of PGA scaling values. In this respect, Table 1 reports normalized restoring force peaks for the selected PGA levels while Figure 4 shows the hysteretic responses of the PS of the 2-DOF hybrid model for the three different levels of excitation.

Table 1 - Normalized peak restoring force.

	PGA $[m/s^2]$		
	1.00	5.00	10.00

$\max_t \frac{\lvert r(t)\rvert}{r_y}$	0.114	0.477	0.774

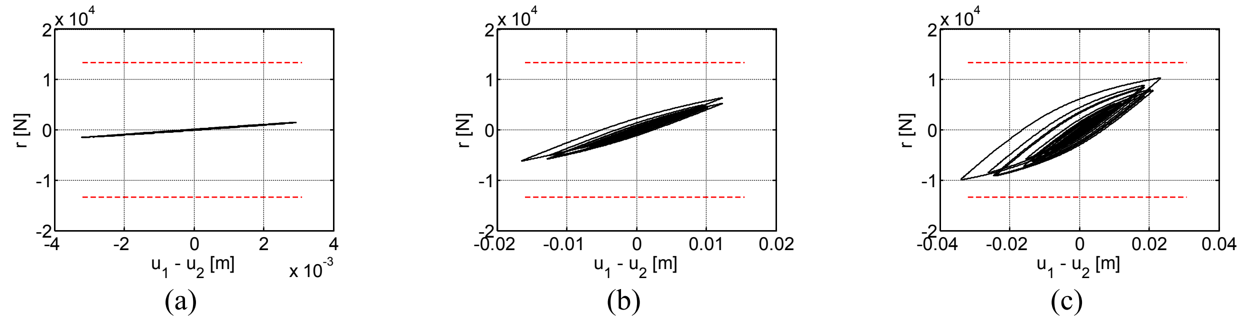


Fig. 4 - Hysteretic responses of the PS of the 2-DOF hybrid model with mean parameter values for: a) $\text{PGA}_1 = 1.00\ m/s^2$; b) $\text{PGA}_2 = 5.00\ m/s^2$; c) $\text{PGA}_3 = 10.00\ m/s^2$. Red dashed lines indicate the inelastic force limits r_y.

Dynamics responses depicted in Figure 4 were computed by solving Eq. (15) considering the reference values of the system parameters using the *ode15s* Matlab solver.

4.2 Reliability analysis setting

The 2-DOF system is assumed to fail when the elongation of the inelastic spring (the PS) exceeds a limit. Thus the model $\mathcal{M}(\boldsymbol{X})$ is defined to compute this maximum absolute elongation as a function of the 2-DOF system mass displacements obtained via HS:

$$\mathcal{M}(\boldsymbol{X}) = \max_t \lvert u_2(\boldsymbol{X},t) - u_1(\boldsymbol{X},t)\rvert \tag{18}$$

where the random vector $\boldsymbol{X}$ gathers the NS spring stiffness parameters k_1 and k_3. Each random parameter was assumed to have a uniform distribution with predefined values of the finite support, mean and variance, as summarized in Table 2.

Table 2: Definition of random input parameters (linear elastic spring stiffnesses).

	$\min[X]$	$\max[X]$	$\mu[X]$	$\sigma[X]$	$COV[X]$
$k_1(N/m)$	2e5	6e5	4e5	1.155e5	0.289
$k_3(N/m)$	5e5	1.5e6	1e6	2.887e5	0.289

As reported in Table 2, mean values of distributions coincide with corresponding reference parameters. A two-dimensional Sobol sequence was used to generate 20,000 pseudo-random samples of the input variables. For each excitation level, the admissible displacement threshold y_{adm} was determined according to Eq. (2-3) via MCS by selecting a target probability of failure, as shown in Table 3. As a result, the same dataset was used to benchmark the performance of the AK-HS reliability method in estimating both failure probabilities and quantiles.

Table 3 - Admissible displacements y_{adm} for different PGA levels and selected probabilities of failure.

		Probability of failure				
		$p_f = 0.20$	$p_f = 0.10$	$p_f = 0.05$	$p_f = 0.02$	$p_f = 0.01$
PGA $[m/s^2]$	1	0.0051	0.0057	0.0059	0.0062	0.0063
	5	0.0254	0.0277	0.0287	0.0300	0.0311
	10	0.0521	0.0576	0.0645	0.0698	0.0719

Table 4 summarizes the accuracy of the failure probability estimates, which is measured using the theoretical coefficient of variation for the 20,000 Monte Carlo samplings, shown in Eq. (19).

$$COV[\hat{p}_f] = \sqrt{\frac{1-\hat{p}_f}{n_{MC}\cdot \hat{p}_f}} \tag{19}$$

Table 4 - Accuracy of failure probability estimates.

	Probability of failure				
	$p_f = 0.20$	$p_f = 0.10$	$p_f = 0.05$	$p_f = 0.02$	$p_f = 0.01$
$COV[\hat{p}_f]$	0.014	0.021	0.031	0.049	0.070

4.3 Reliability analysis results

This subsection discusses the results of the numerical validation of the AK-HS reliability method. In this context and in this study, a single numerical solution of Eq. (16) replaces a single HS run that involves a new PS. With this in mind, the number of HSs must be as small as possible because each run involves a test of the PS in the laboratory and, thus, adds significantly to the costs and effort. Therefore, the size of the starting ED was set to 12 samples while its maximum size was limited to 32 samples, i.e. 32 HSs. According to previous studies [14] conducted on the same benchmark 2-DOF hybrid system and focusing on Uncertainty Propagation (UP) and Global Sensitivity Analysis (GSA) based on Polynomial Chaos Expansion (PCE), this was sufficient to provide accurate estimates of statistical moments and Sobol' indices of output quantities in the case of three input random parameters. The Kriging metamodels were estimated using the UQLab software framework developed by the Chair of Risk, Safety and Uncertainty Quantification in ETH Zurich [15] using the *Structural Reliability* module [16]. Figure 5 provides an overview of the learning process underlying the AK-HS reliability method for the case with $PGA = 10\,m/s^2$ and $p_f = 10\%$.

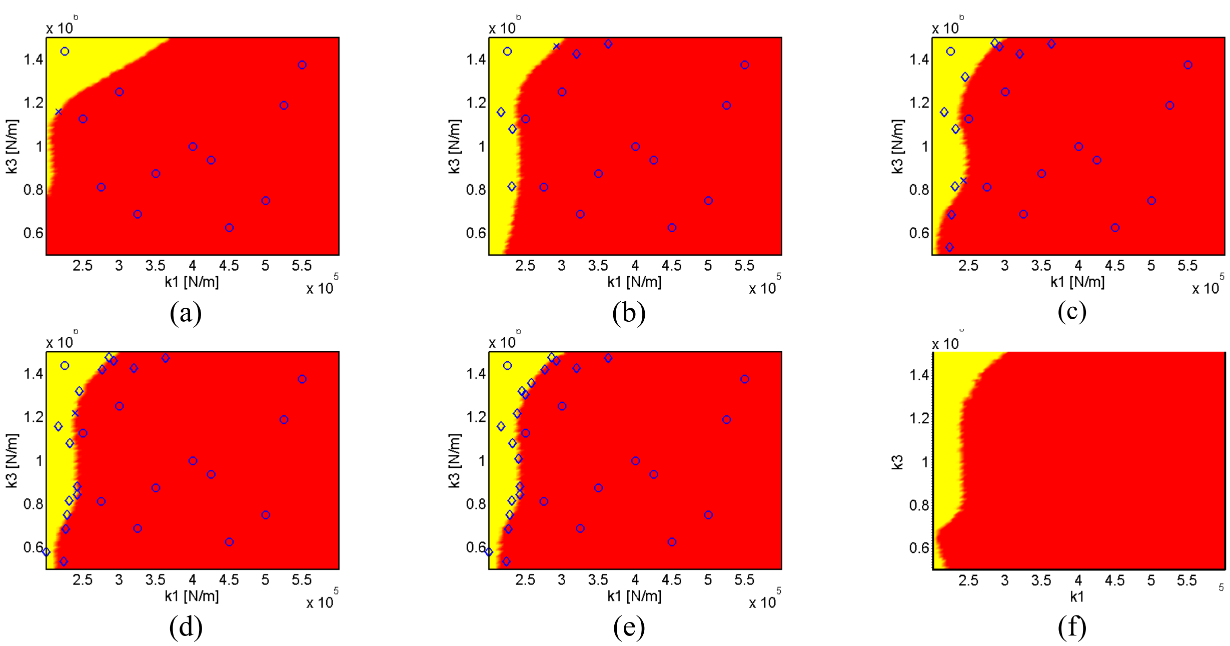


Fig. 5 - Overview of the failure domain estimation process for the case of $PGA = 10\,m/s^2$ and $p_f = 0.10$ based on: (a) 12 samples (initial training dataset); (b) 17 samples; (c) 22 samples; (d) 27 samples; (e) 32 samples; and compared to the reference MCS (f). The estimated failure domain is depicted in yellow; the circular dots represent the initial ED set, the diamonds represents the additional samples selected by the learning function, and the cross dots indicate the best suited sample for the next HS.

As seen in Figure 5a, the initial pool of 12 samples outlines a rough approximation of the actual failure domain, which is refined by the learning process (Figures 5, b-c-d-e). In detail, in each iteration, the learning function locates the next best-suited sample, represented by a cross dot, on the estimated boundary of the failure domain, where the probability of misclassification is the highest. After 20 iterations, the AK-HS reliability method provides a good approximation of the reference failure domain calculated via MCS as depicted in Figure 5f. In order to quantify the performance of the AK-HS reliability method, the following scores are introduced:

- $\rho_{\hat{p}_f} = \frac{|p_f - \hat{p}_f|}{p_f}$: Normalized Error of failure Probability Estimate (NEPE);
- $\rho_{\hat{q}_\alpha} = \frac{|q_\alpha - \hat{q}_\alpha|}{std(y)}$: Normalized Error of Quantile Estimate (NEQE);
- $\varepsilon_{\hat{p}_f} = \frac{\hat{p}_f^+ - \hat{p}_f^-}{\hat{p}_f}$: Normalized Dispersion of failure Probability Estimate (NDPE);
- $\varepsilon_{\hat{q}_\alpha} = \frac{\hat{q}_\alpha^+ - \hat{q}_\alpha^-}{std(y)}$: Normalized Dispersion of Quantile Estimate (NDQE);

where $std(y)$ represents the standard deviation of the 2-DOF hybrid model response with respect to the overall input probability space. All scores refer to final estimates, that is, after 32 HSs. Tables 5a and 5b summarize NEPE and NEQE scores for the investigated values of failure probability and excitation levels. Analogously, Table 6a and 6b summarizes NDPE and NDQE scores.

Table 5a - NEPE $\rho_{\hat{p}_f}$ scores.

		Probability of failure p_f				
		0.20	0.10	0.05	0.02	0.01
PGA $[m/s^2]$	1	0.105	0.031	0.051	0.017	0.040
	5	0.010	0.038	0.036	0.017	0.184
	10	0.012	0.050	0.016	0.030	0.080

Table 5b - NEQE $\rho_{\hat{q}_\alpha}$ scores.

		Probability of failure p_f				
		0.20	0.10	0.05	0.02	0.01
PGA $[m/s^2]$	1	0.082	0.012	0.006	0.002	0.002
	5	0.007	0.010	0.007	0.004	0.036
	10	0.006	0.024	0.008	0.006	0.012

Table 6a - NDPE $\varepsilon_{\hat{p}_f}$scores.

		Probability of failure p_f				
		0.20	0.10	0.05	0.02	0.01
PGA $[m/s^2]$	1	0.466	1.269	0.289	2.317	8.468
	5	1.027	0.803	0.616	1.681	63.542
	10	1.185	0.501	2.078	0.312	13.517

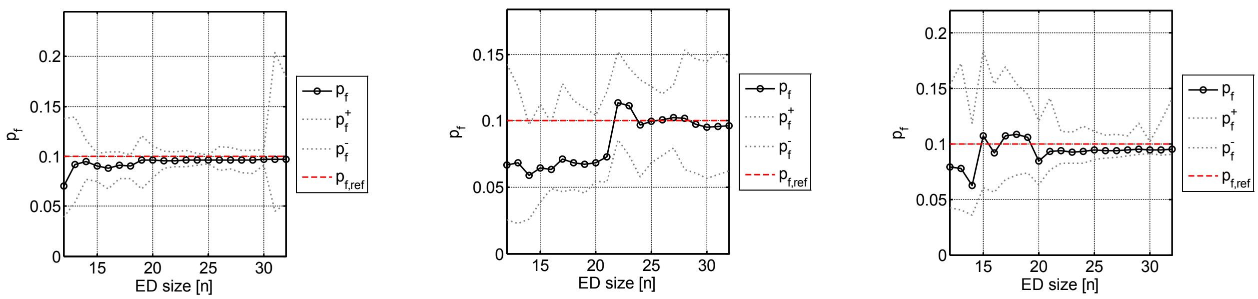
Table 6b - NDQE $\varepsilon_{\hat{q}_\alpha}$scores.

		Probability of failure p_f				
		0.20	0.10	0.05	0.02	0.01
PGA $[m/s^2]$	1	0.300	0.214	0.026	0.186	0.361
	5	0.354	0.138	0.074	0.152	5.959
	10	0.462	0.187	0.529	0.050	1.213

In particular, NEPE and NEQE scores highlight the presence of bias in statistic estimates. The results shown in Table 5a and 5b are quite promising: NEPE and NEQE scores indicate that failure probability and quantile estimates, respectively, have small errors and are almost un-biased for all cases. It is important to remember that investigated quantiles and failure probabilities are complementary, that is, $q_\alpha = y_{\mathrm{adm}}$ and $\alpha = p_f$. As a result, the same dataset and Kriging metamodels were used to benchmark the AK-HS reliability method in providing estimates of both the probability of failure and the quantiles.

Dispersion-related NDPE and NDQE scores indicate the width of the confidence intervals. As reported in Table 6a and 6b, the width of confidence bounds remains relatively large for both failure probabilities and quantiles. In addition, the two scores do not show any stable trend with respect to the explored grid of parameters.

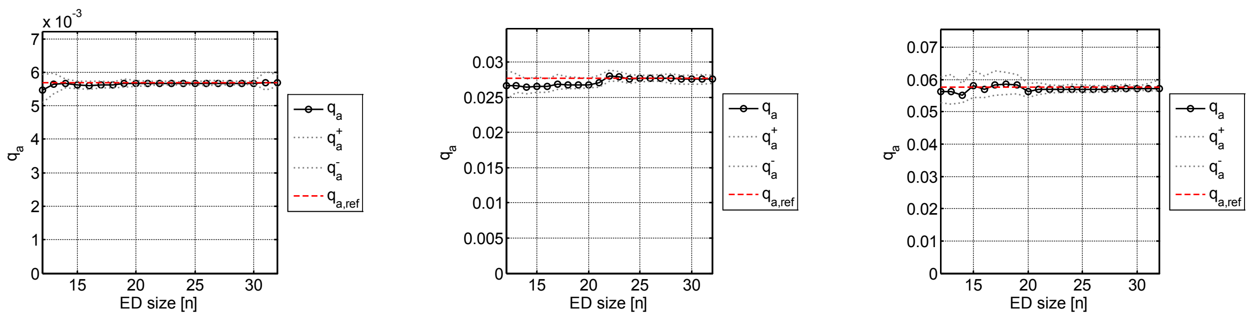
A careful reader can notice that all scores tend to show higher values in the case of $PGA = 5\,m/s^2$ and $p_f = 0.01$. In our opinion, such outliers potentially hide a threshold behavior of the system response with respect to input parameters. In particular, this excitation level is supposed to produce a slightly nonlinear system response when mean values of parameters are selected. A more pronounced either linear or nonlinear behavior could be observed in different regions of the parameter space, thus making the overall limit state surface sharper owing to transition zones. Figure 6 offers a close-up view of the trend of the failure probability estimates in the case of $p_f =$ 10%.



(a) (b) (c)

Fig. 6 - Trends of failure probability estimates in the case of $p_f = 0.10$ and: (a) $PGA = 1\,m/s^2$; (b) $PGA = 5\,m/s^2$; (c) $PGA = 10\,m/s^2$.

As shown in Figure 6, convergence is achieved with less than 32 HSs in all cases. However, the confidence bounds keep oscillating without any evident trend. Analogously, Figure 7 offers a close-up view of the trend of quantile estimates in the case of $p_f = 10\%$.



(a) (b) (c)

Fig. 7 - Trends of quantile estimates in the case of $p_f = 0.10$ and: (a) $PGA = 1\,m/s^2$; (b) $PGA = 5\,m/s^2$; (c) $PGA = 10\,m/s^2$.

In this particular case, apparently, the starting ED is sufficient to provide reliable estimates of the quantile and the confidence bounds are narrows. Even though this is, most likely, a fortuitous occurrence, such result corroborates the effectiveness of the proposed AK-HS reliability method.

5 Conclusions

Reliability analysis aims to determine the probability of failure of a stochastic system and typically relies on purely numerical simulations. Nevertheless, thrusting mathematical models up to failure can be questioned, especially for structural components characterized by strongly nonlinear response. In this paper, hybrid simulation is presented as suitable alternative to evaluate the structural system response for the purpose of RA, thus reducing dramatically the amount of epistemic uncertainty. Since HS suffers the same limitations as the large FE models, the proposed procedure relies on non-intrusive Kriging metamodeling. An initial pool of samples is selected to cover the entire probability space of input random parameters and the corresponding initial HS provide the model response. A learning function dictates the sample parameters for the following HSs to minimize the uncertainty in the localization of the failure domain, which is predicted by a Kriging metamodel. Probability of failure and quantile estimates are calculated using the last Kriging metamodel via MCS. The benchmark validation of the resulting AK-HS reliability method is presented for a 2-DOF spring series system with one inelastic spring modeled as the PS and the other two elastic springs modeled as the NSs of a hybrid model. The benchmark example proves that fewer than 32 HSs are sufficient to accurately model the reliability of systems in this class of complexity with target failure probabilities as low as 1%. While still large, this number of experiments is not uncommon in quasi-static cyclic test campaigns aimed at characterizing response of structural elements [17]. Future work is aimed at incorporating RA methods into guidelines for conducting HS experimental campaigns.

6 References

[1] Lemaire, Maurice. Structural reliability. John Wiley & Sons, 2013.

[2] Kiureghian, A. Der, & Ditlevsen, O.. Aleatory or epistemic? Does it matter? Structural Safety, 31(2), 105–112. doi:10.1016/j.strusafe.2008.06.020, 2009.

[3] Takanashi, K., Udagawa, K., Seki, M., Okada, T. and Tanaka, H., "Non-linear earthquake response analysis of structures by a computer-actuator on-line system", Bulletin of Earthquake Resistant Structure Research Center, Institute of Industrial Science, University of Tokyo, Tokyo, 1975.

[4] Stojadinovic, B., G. Mosqueda, and S. A. Mahin.. ''Event-Driven Control System for Geographically Distributed Hybrid Simulation'', ASCE Journal of Structural Engineering, Vol. 132, No. 1, pp. 68-77, 2006.

[5] Bursi O.S., Wagg D.J.E.. Modern Testing Techniques for Structural Systems, CISM International Centre for Mechanical Sciences, vol 502. Springer, Vienna, 2008.

[6] Berveiller, M., Sudret, B., & Lemaire, M.. Stochastic finite element: a non intrusive approach by regression. Revue Européenne de Mécanique Numérique, 15(1-2-3), 81–92. doi:10.3166/remn.15.81-92, 2006.

[7] Echard, B., Gayton, N., & Lemaire, M.. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation. Structural Safety, 33(2), 145–154. doi:10.1016/j.strusafe.2011.01.002, 2011.

[8] Schöbi, R., Sudret, B., & Marelli, S.. Rare Event Estimation Using Polynomial-Chaos Kriging. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering, D4016002. doi:10.1061/AJRUA6.0000870, 2016.

[9] Sobol, I.M., "Distribution of points in a cube and approximate evaluation of integrals". Zh. Vych. Mat. Mat. Fiz. 7: 784–802 (in Russian); U.S.S.R Comput. Maths. Math. Phys. 7: 86–112 (in English), 1967.

[10] Santner, Thomas J., Brian J. Williams, and William I. Notz. The design and analysis of computer experiments. Springer Science & Business Media, 2013.

[11] Murphy, Kevin P. Machine learning: a probabilistic perspective. MIT press, 2012.

[12] Ismail, Mohammed, Fayçal Ikhouane, and José Rodellar. "The hysteresis Bouc-Wen model, a survey." Archives of Computational Methods in Engineering 16.2: 161-188, 2009.

[13] PEER. Pacific Earthquake Engineering Research (PEER) Center, Ground Motion Database. <http://ngawest2.berkeley.edu/site>, May 24, 2016.

[14] Marelli, S. & Sudret, B. UQLab: A Framework for Uncertainty Quantification in Matlab 257 Vulnerability, Uncertainty, and Risk (Proc. 2nd Int. Conf. on Vulnerability, Risk Analysis and Management (ICVRAM2014), Liverpool, United Kingdom), 2554-2563, 2014.

[15] C. Lataniotis, S. Marelli and B. Sudret, UQLab user manual – Kriging (Gaussian process modelling), Report UQLab-V0.9-105, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich (2015).

[16] Abbiati G., Marelli S., Bursi O.S., Sudret B., Stojadinovic B., Uncertainty Propagation and Global Sensitivity Analysis in Hybrid Simulation Using Polynomial Chaos Expansion. The Fifteenth International Conference on Civil, Structural and Environmental Engineering Computing CIVIL-SOFT-COMP (CSC2015) Prague, Czech Republic 1-4 September 2015.

[17] Mahin S., Malley J., Hamburger R., Mahoney M., Overview of a Program for Reduction of Earthquake Hazards in Steel Frame Structures, Proceedings of the 12th World Conference on Earthquake Engineering, paper #2541, Auckland, New Zealand, 2000.