SECOND EUROPEAN CONFERENCE ON EARTHQUAKE ENGINEERING
AND SEISMOLOGY, ISTANBUL AUG. 25-29, 2014

SURROGATE MODELING IMPLEMENTATION FOR ASSESSMENT OF SEISMIC RISK UTILIZING STOCHASTIC GROUND MOTION MODELING

Ioannis GIDARIS[1], Alexandros TAFLANIDIS[2] and George MAVROEIDIS[3]

ABSTRACT

An efficient computational framework is presented for seismic risk assessment within a probabilistic modelling approach that utilizes stochastic ground motion models for the seismic hazard characterization. The framework is based on a kriging surrogate modelling approach for efficiently approximating the structural response with respect to the structural and ground motion parameters that are considered as uncertain, whereas the stochastic characteristics of the seismic excitation are addressed by assuming that under the influence of white noise each response quantity follows a lognormal distribution. The surrogate model is then used for approximating the structural response when addressing the uncertainty in the various model parameters and ultimately estimating seismic risk. Seismic hazard descriptions including or not near-fault pulses are separately investigated, whereas the impact of the model prediction error stemming from the kriging metamodel is directly incorporated within the seismic risk quantification and assessment. The proposed framework is demonstrated for risk assessment of a benchmark reinforced concrete office building.

INTRODUCTION

Evaluation of seismic risk for structural system requires (i) adoption of appropriate models for the structural system itself and for the earthquake excitation and (ii) quantification of the uncertainties involved in these models. Undoubtedly the most critical component in this approach is the appropriate description of the seismic excitation, and for applications involving dynamic analysis, characterization of the entire ground motion history is required for this purpose. The growing interest in performance-based earthquake engineering (PBEE) in the last decade has intensified this need. Within PBEE seismic risk is described by addressing the entire spectrum of structural response, ranging from linear to grossly nonlinear, up to structural collapse, and this approach requires a realistic characterization of the entire earthquake acceleration time-history. A modelling approach to establish this, gaining increasing interest within the structural engineering community, is the use of stochastic ground motion models. These models are based on modulation of a white noise sequence through functions that address spectral and temporal characteristics of the excitation. The parameters of these functions (such as duration of excitation or parameters related to frequency content) can be related to seismicity (moment magnitude and rupture distance) and site characteristics (soil conditions) by appropriate predictive relationships. Description of the uncertainty in these predictive relationships and in the regional seismicity leads then to a complete probabilistic model for describing seismic hazard in terms

[1] Ph.D Candidate, University of Notre Dame, Notre Dame,IN, igidaris@nd.edu
[2] Associate Professor, University of Notre Dame, Notre Dame,IN, a.taflanidis@nd.edu
[3] Assistant Professor, University of Notre Dame, Notre Dame,IN, g.mavroeidis@nd.edu

of acceleration time-histories. Coupled with a probabilistic description of the uncertainties in the structural characteristics, this approach ultimately facilitates a comprehensive quantification of seismic risk. Evaluation of this risk can be then performed through stochastic simulation (Taflanidis and Beck 2009) which can facilitate high-accuracy estimates for the seismic risk, though at a large computational burden, especially for applications involving complex, nonlinear models.

This study offers an efficient computational framework, relying on surrogate modelling, to alleviate this burden and facilitate a highly efficient seismic risk estimation within such a modelling context, i.e. when stochastic ground motion models are utilized for the seismic hazard description. Kriging is adopted in particular as a surrogate model for this purpose whereas for the structural performance the fragility of different engineering demand parameters (EDPs) is considered. The assessment of seismic risk for a four-story reinforced concrete office building is considered as an illustrative example for the proposed framework, whereas a point-source stochastic ground motion model is adopted to describe acceleration time-histories. Additionally, the occurrence of "forward directivity" pulses is considered in the description of the seismic hazard which can be very important implications for the structural response in locations close to active faults (Mavroeidis, Dong et al. 2004). In the context of this example different interesting aspects related to seismic risk are investigated. In particular, (a) how does inclusion of the "forward-directivity" effect, which produces the pulse-like ground motions, in the seismic hazard description affect the accuracy of the surrogate modeling and additionaly the seismic risk and (b) how can the prediction error related to the kriging approximation be incorporated within the seismic risk quantification and assessment?

SEISMIC RISK QUANTIFICATION

Evaluation of seismic risk is facilitated through the approach shown in Figure 1, by augmentation of appropriate models for the structural system, the earthquake excitation and the system performance. The combination of the first two models provides the structural response, denoted **z** herein, and in the approach adopted here this is established in terms of nonlinear time-history analysis and ultimately approximated through a kriging surrogate model. The performance evaluation model assesses, then, the favorability of this response **z**, based on chosen criteria of the system's stakeholders. In particular a stochastic ground motion model is adopted here as an excitation model which entails, as presented in detail in the following sections, modulation of a stochastic sequence $\mathbf{w} \in W$ through appropriate time and frequency domain functions.

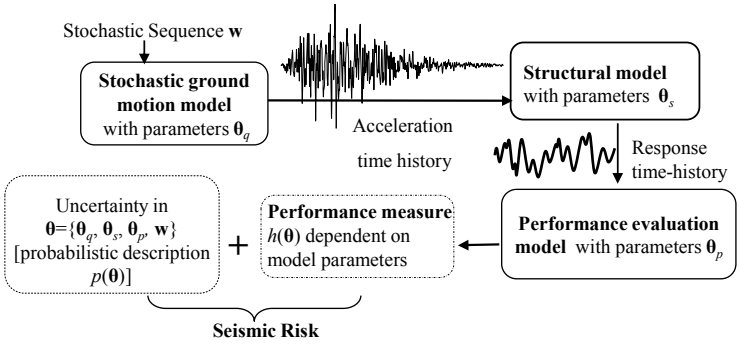

Figure 1. Augmented model for seismic risk quantification

The uncertainty in the characteristics of the model in Figure 1 and in the stochastic sequence **w** can be described by assigning a probability model to them. In this context, let $\boldsymbol{\theta} \in \Theta$, denote the n_θ-dimensional augmented vector of model parameters, where Θ represents the space of possible model parameter values. Vector $\boldsymbol{\theta}$ is composed of all the model parameters for the individual structural system, excitation, and performance evaluation models as indicated in Figure 1. For addressing the uncertainty in $\boldsymbol{\theta}$, a probability density function (PDF) $p(\boldsymbol{\theta})$, is assigned to it based on our available knowledge. Let, also the performance of the structural system, for specific $\boldsymbol{\theta}$, be characterized by the

risk consequence measure $h(\boldsymbol{\theta},\mathbf{w})$, ultimately estimated based on the predicted response $\mathbf{z}$. Seismic risk H, is then given by the expected value of $h(\boldsymbol{\theta},\mathbf{w})$, over the probability models:

$$H = \int_{\Theta}\int_{W} h(\boldsymbol{\theta},\mathbf{w})p(\mathbf{w})p(\boldsymbol{\theta})d\mathbf{w}d\boldsymbol{\theta} \tag{1}$$

Through appropriate selection of the risk consequence measure $h(\boldsymbol{\theta})$ different risk quantifications can be addressed within this framework. In this study the fragility of different EDPs is considered as risk quantification. This corresponds to a popular consequence measure appearing in different seismic risk applications, for example within structural reliability or life-cycle cost estimation (Ellingwood 2001; Goulet, Haselton et al. 2007; Taflanidis and Beck 2009; Taflanidis and Jia 2011), and is defined as the probability that an EDP, corresponding to some response quantity z_j (e.g. peak interstory drift, peak floor acceleration, etc.) will exceeds some threshold β that determines acceptable structural performance. It is common to incorporate a prediction error in this definition (Porter, Kennedy et al. 2006; Taflanidis, Vetter et al. 2013); this can be equivalently considered as the threshold corresponding to an uncertain quantity ($\beta\cdot\varepsilon_\beta$) with ε_β having a lognormal distribution with median equal to one and logarithmic standard deviation σ_β. This then leads to

$$h(\boldsymbol{\theta},\mathbf{w}) = P[z_j > \beta\varepsilon_\beta] = P[\varepsilon_\beta \le z_j/\beta] = P[\ln(\varepsilon_\beta) \le \ln(z_j) - \ln(\beta)] = \Phi\left(\frac{\ln(z_j)-\ln(\beta)}{\sigma_\beta}\right) \tag{2}$$

where $\Phi(.)$ denotes the standard normal cumulative distribution. This equation (2) corresponds to common definition of fragility function in earthquake engineering and for σ_β=0, corresponding to no uncertainty in the description of β it simplifies to the indicator function, which is one if $z_j>\varepsilon_\beta$ and zero if not.

STOCHASTIC GROUND MOTION MODELING FOR SEISMIC HAZARD DESCRIPTION

For quantifying seismic hazard a stochastic ground motion model is considered in this study as depicted earlier in Figure 1, whereas for a comprehensive characterization of this hazard in regions close to active faults, forward-directivity pulse effects are additionally incorporated. For this purpose, the high-frequency (broadband) and long period (pulse) components of the motion are independently modeled and then combined to form the acceleration time history. For cases where the former only component is needed (no near-fault effects) the long period pulse component is simply ignored.

High frequency (broadband) component: For modeling the higher-frequency (>0.1-0.2 Hz) component of ground motions, a point source stochastic model is chosen here. This model is based on a parametric description of the ground motion's radiation spectrum, dependent on the earthquake characteristics (earthquake magnitude, M, and rupture distance, r –defined as closest distance to the fault plane) and soil profile. This spectrum incorporates the influence of the source, path and local site conditions. The duration of the ground motion is addressed through an envelope function, which again depends on M and r. These frequency and time domain functions completely characterize the model and their properties are provided by predictive relationships that relate them directly to the seismic hazard, i.e. to M and r. The same characteristics for this model as in (Taflanidis and Beck 2009) are used here. Particularly, the two-corner point-source model described in detail (Atkinson and Silva 2000) is selected for the source spectrum because of its partial compatibility with finite fault models. This compatibility is important because we are also interested in ground motions in close proximity to active faults, for which the assumptions behind point-source models are challenged (Atkinson and Silva 2000). The high-frequency component of the time history, for specific values of M and r, is ultimately developed according to this model by modifying a white-noise sequence first by the time domain envelop function and subsequently by the frequency-domain radiation spectrum.

Long period pulse component: For characterizing the pulse, a simple analytical model (Mavroeidis and Papageorgiou 2003) is selected, describing the pulse through the following velocity time-history

$$V = \frac{A_p}{2}\left[1+\cos\left(\frac{2\pi}{\gamma_p}\frac{(t-t_o)}{T_p}\right)\right]\cos\left[2\pi\frac{(t-t_o)}{T_p}+\nu_p\right], \quad t\in\left[t_o-\frac{\gamma_p T_p}{2}, t_o+\frac{\gamma_p T_p}{2}\right]; =0 \text{ otherwise} \quad (3)$$

where A_p, T_p, γ_p, ν_p and t_o describe the signal amplitude, pulse period, oscillatory character (i.e. number of half cycles), phase angle, and time shift to specify the envelope's peak, respectively. The first two pulse characteristics can estimated by predictive relationships (Mavroeidis and Papageorgiou 2003; Halldórsson, Mavroeidis et al. 2010) relating them to M and r

$$\log_{10}(A_p/0.9) = 2.04 - 0.032r + e_A \quad (4)$$

$$\log_{10}(T_p) = -2.9 + 0.5M + e_T \quad (5)$$

with e_A and e_T zero mean Gaussian variables with standard deviations 0.187 and 0.143, respectively. The other two γ_p and ν_p can be considered as independent parameters, whereas t_o can be selected based on the anticipated peak of the excitation.

Final ground motion: These two components are ultimately combined to yield a near-fault ground motion time history by the procedure proposed in (Mavroeidis and Papageorgiou 2003). The resultant stochastic ground motion model has as model parameters the moment magnitude M, the rupture distance r and the pulse characteristics A, f_p, γ_p and ν_p, whereas t_o is defined through the process of superimposing the two components.

KRIGING SURROGATE MODEL FORMULATION FOR SEISMIC RISK ASSESSMENT

For efficiently estimating seismic risk a kriging metamodel is developed here to approximate the structural response with respect to the structural $\boldsymbol{\theta}_s$ and excitation $\boldsymbol{\theta}_g$ uncertain parameters, whereas the stochastic characteristics of the ground motion are addressed by assuming that under the influence of the white noise ($\mathbf{w}$ in Figure 1) each response quantity z_j follows a lognormal distribution with median $\bar{z}_j$ and coefficient of variation σ_{z_j}, which corresponds to a common assumption within earthquake engineering (Zhang and Foschi 2004; Jalayer and Cornell 2009). This means that $z_j = \varepsilon_j \bar{z}_j$ with ε_j having a lognormal distribution with median equal to one and logarithmic standard deviation σ_{z_j}. Under these assumptions, approximately addressing the impact of the white noise $\mathbf{w}$, the risk integral in (1) simplifies to

$$H = \int_{\Theta} h(\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (6)$$

with consequence measure given by

$$h(\boldsymbol{\theta}) = P[z_j \geq \beta\varepsilon_\beta] = P[\ln(\varepsilon_\beta) \leq \ln(z_j) - \ln(\beta)] = P[\ln(\varepsilon_\beta) \leq \ln(\bar{z}_j) + \ln(\varepsilon_j) - \ln(\beta)]$$
$$= P[\ln(\varepsilon_\beta) - \ln(\varepsilon_j) \leq \ln(\bar{z}_j) - \ln(\beta)] = \Phi\left(\frac{\ln(\bar{z}_j) - \ln(\beta_i)}{\sqrt{\sigma_\beta^2 + \sigma_{z_j}^2}}\right) \quad (7)$$

where the last equality is based on the fact that since $\ln(\varepsilon_\beta)$ and $\ln(\varepsilon_j)$ are zero mean Normal variables their difference will be also a Normal variable with zero mean and standard deviation $\sqrt{\sigma_\beta^2 + \sigma_{z_j}^2}$ (Ditlevsen and Madsen 1996).

The kriging metamodel is ultimately formulated through the procedure shown in Figure 2 to provide predictions for the statistical quantities needed for evaluation of (7), corresponding to the statistics for the EDPs of interest due to the influence of the white noise (i.e. $\ln(\bar{z}_j)$ and σ_{z_j}). Let $\mathbf{y}$ denote the n_y dimensional vector of such quantities to be approximated by the metamodel, with y_k denoting the k^{th} component, and $\mathbf{x} = [\boldsymbol{\theta}_g \ \boldsymbol{\theta}_s]$ the n_x dimensional augmented vector of the uncertain parameters for the kriging formulation. Note that this vector does not need to include the performance evaluation model parameters, since the surrogate model approximates simply the structural response and not the risk consequence measure directly. A separate metamodel is formed for the case that the ground motion model incorporates the near-fault pulse effects and when only the broadband far-field part of the excitation is taken into account. These two seismic hazard descriptions are abbreviated herein as P and NP, respectively. For forming the metamodel initially, a database with *n observations* is obtained that provides information for the **x-y** pair. For this purpose n samples for $\{\mathbf{x}^i \ i=1,\ldots,n\}$, also known as *support points*, are created following a latin hybercube grid over the expected range of values possible for each x_i. This is ultimately informed by the range Θ that will be needed in the evaluation of (7). It should be stressed that this does not require a firm definition for $p(\boldsymbol{\theta})$, simply a knowledge to the range for which the kriging metamodel will be used so that the support points extend over this range (this allows to ultimately avoid extrapolations). Stochastic ground motions are then generated according to the excitation model and the structural response is numerically evaluated. The influence of the white noise is addressed by considering in this study $n_w = 100$ different samples for it (for each $\mathbf{x}^i$) and the statistics under these samples ultimately quantify the response sample $\mathbf{y}^i$ (corresponding to $\ln(\bar{z}_j)$ and σ_{z_j} for all EDPs).

Using this dataset the kriging model is then obtained, providing ultimately approximation $y_k = \hat{y}_k + \varepsilon_k$, where $\hat{y}_k$ stands for the mean and whereas ε_k is a Gaussian variable with zero mean and standard deviation σ_k (Lophaven 2002). The fundamental building blocks of kriging are the n_p dimensional basis vector, $\mathbf{f}(\mathbf{x})$, and the correlation function, $R(\mathbf{x}^l,\mathbf{x}^m)$, defined through parameter vector $\mathbf{s}$. Typical selections are made for both here, for the former a full quadratic basis and for the latter a generalized exponential correlation, leading to

$$\mathbf{f}(\mathbf{x}) = [1 \ \ x_1 \ \cdots \ x_{n_x} \ \ x_1^2 \ x_1x_2 \ \cdots \ x_{n_x}^2] \ ; n_p = (n_x+1)(n_x+2)/2 \tag{8}$$

$$R(\mathbf{x}^l,\mathbf{x}^m) = \prod_{i=1}^{n_x} \exp[-s_i \,|\, \mathbf{x}_i^l - \mathbf{x}_i^m \,|^{s_{n_x+1}}] \ ; \mathbf{s} = [s_1 \ \cdots \ s_{n_x+1}] \tag{9}$$

Then for the set of n observations (training set) with input matrix $\mathbf{X} = [\mathbf{x}^1 \ \cdots \ \mathbf{x}^n]^T$ and corresponding output $\mathbf{Y} = [\mathbf{y}(\mathbf{x}^1) \ \cdots \ \mathbf{y}(\mathbf{x}^n)]^T$, we define the basis matrix $\mathbf{F}=[f(\mathbf{x}^1) \ \ldots \ f(\mathbf{x}^n)]^T$ and the correlation matrix $\mathbf{R}$ with the lm-element defined as $R(\mathbf{x}^l,\mathbf{x}^m)$, $l,m=1, \ \ldots, \ n$. Also for every new input $\mathbf{x}$, we define the correlation vector $\mathbf{r}(\mathbf{x}) = [R(\mathbf{x},\mathbf{x}^1) \ \ldots \ R(\mathbf{x},\mathbf{x}^n)]^T$ between the input and each of the elements of $\mathbf{X}$. Then the mean kriging approximation is given by (Lophaven 2002)

$$\begin{aligned} &\hat{\mathbf{y}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T\boldsymbol{\alpha}^* + \mathbf{r}(\mathbf{x})^T\boldsymbol{\beta}^* \\ &\boldsymbol{\alpha}^* = (\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{R}^{-1}\mathbf{Y}; \ \boldsymbol{\beta}^* = \mathbf{R}^{-1}(\mathbf{Y}-\mathbf{F}\boldsymbol{\alpha}^*) \end{aligned} \tag{10}$$

Through the proper tuning of the parameters $\mathbf{s}$ of the correlation function, kriging can efficiently approximate very complex functions. The optimal selection of $\mathbf{s}$ is based on the Maximum Likelihood Estimation (MLE) principle, where the likelihood is defined as the probability of the n observations, and maximizing this likelihood with respect to $\mathbf{s}$ ultimately corresponds to the optimization problem (Lophaven 2002)

$$\mathbf{s}^* = \arg\min_{\mathbf{s}} \left[|\mathbf{R}|^{\frac{1}{n}} \sum_{k=1}^{n} \tilde{\sigma}_k^2 \right] \tag{11}$$

where |.| stands for determinant of a matrix and $\tilde{\sigma}_k^2$, k = 1,.., n correspond to the diagonal elements of matrix $(\mathbf{Y}-\mathbf{F}\boldsymbol{\alpha}^*)^T\mathbf{R}^{-1}(\mathbf{Y}-\mathbf{F}\boldsymbol{\alpha}^*)/n$. The prediction error variance $\sigma_k^2(\mathbf{x})$ for y_k and input $\mathbf{x}$ is given by

$$\sigma_k^2(\mathbf{x}) = \tilde{\sigma}_k^2[1+\mathbf{u}^T(\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{u}-\mathbf{r}(\mathbf{x})^T\mathbf{R}^{-1}\mathbf{r}(\mathbf{x})];\ \mathbf{u}=\mathbf{F}^T\mathbf{R}^{-1}\mathbf{r}(\mathbf{x})-\mathbf{f}(\mathbf{x}) \tag{12}$$

This local estimate (i.e. dependent on **x** values) for the kriging prediction error statistics can be additionally explicitly incorporated within the risk assessment. For the approximation to the mean log predictions $\ln(\bar{z}_j)=\ln(\hat{\bar{z}}_j)+\varepsilon_{\bar{z}_j}$, where the standard deviation for $\varepsilon_{\bar{z}_j}$ is $\sigma_{\varepsilon_{\bar{z}_j}}$, this leads, following a similar approach as in (7), to the following transformation for the risk consequence measure

$$\begin{aligned} h(\boldsymbol{\theta}) &= P[\ln(\varepsilon_\beta)-\ln(\varepsilon_j)\le\ln(\bar{z}_j)-\ln(\beta)] = P[\ln(\varepsilon_\beta)-\ln(\varepsilon_j)\le\ln(\hat{\bar{z}}_j)+\varepsilon_{\bar{z}_j}-\ln(\beta)]\ln(\bar{z}_j) \\ &= P[\ln(\varepsilon_\beta)-\ln(\varepsilon_j)-\varepsilon_{\bar{z}_j}\le\ln(\hat{\bar{z}}_j)-\ln(\beta)] = \Phi\left(\frac{\ln(\hat{\bar{z}}_j/\beta_i)}{\sqrt{\sigma_\beta^2+\hat{\sigma}_{\varepsilon_{\bar{z}_j}}{}^2+\sigma_{z_j}^2}}\right) \end{aligned} \tag{13}$$

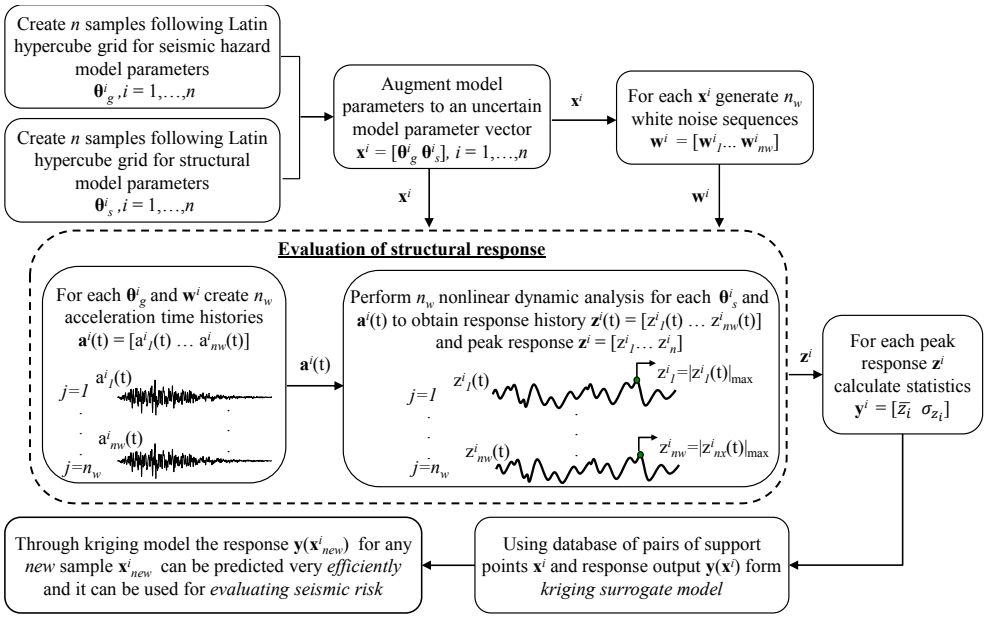

Figure 2. Schematic flowchart of the procedure for creating kriging surrogate model

Ultimately (10) provides a computational efficient approximation to structural response. The computational intensive aspect of the entire formulation is the development of the database **Y** which requires time-history analysis for a large number of model parameters, to populate the entire region for which **x** is anticipated to take values in, and sufficient number of white noise samples, to address the resultant variability in the response. This needs to be performed, though, only once. As soon as the kriging metamodel is established based on this database, it can be then used to efficiently predict the responses for any other **x** desired. The probabilistic integral for calculating seismic risk in (6) can be then estimated using stochastic simulation and utilizing (13), with the structural response provided computationally efficiently through the kriging approximation. Using a finite number, N of samples of $\boldsymbol{\theta}$ drawn from proposal densities $q(\boldsymbol{\theta})$ this estimate for H is given by

$$\hat{H} = \frac{1}{N}\sum_{i_s=1}^{N} h(\boldsymbol{\theta}^{i_s})\frac{p(\boldsymbol{\theta}^{i_s})}{q(\boldsymbol{\theta}^{i_s})} \tag{14}$$

where $\boldsymbol{\theta}^{i_s}$ denotes he sample used in the i_s^{th} simulation. The proposal densities are used to improve efficiency of this estimation, by focusing the computational effort on regions of Θ space that contribute more to the integrand of the integral (6) [or equivalently (1)] (Robert and Casella 2004). Simplest choice (corresponding to Monte Carlo simulation) is to use $q(.)=p(.)$.

ILLUSTRATIVE EXAMPLE

For the illustrative example, the four-story reinforced concrete office building used in (Goulet, Haselton et al. 2007) (corresponds to design A in this reference) and shown in Figure 3 is considered. The structure corresponds to a four-story moment-frame office building designed to comply with the 2003 International Building Code (IBC 2003) and ASCE7-02. The structural layout of the building represents a four bay by six bay plan with 9.10 span lengths. The lateral system consists of two exterior moment resisting frames in each direction, with interior intermediate gravity frames. The analysis here, as in the case in (Goulet, Haselton et al. 2007) is performed for seismic excitation along the shorter building direction. The first storey is 4.6 m height and the three stories above are 4.00 m height each. The column elements size varies from (61 x 71 cm) to (76 x 100 cm) between different stories and the beam elements from (81 x 61 cm) to (107 x 61 cm).

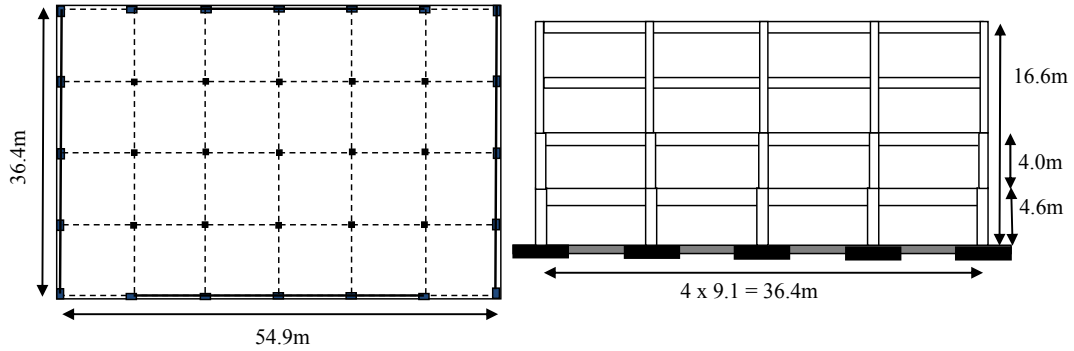

Figure 3. Plan (left) and elevation of the perimeter-frame in the short building direction (right) of the four-story office benchmark building. In the plan view gravity frames are indicated with dashed lines and moment resisting frames with solid lines.

The nonlinear analysis model corresponds to the two dimensional four bay frame, modelled in OpenSees (McKenna 2011). The simulation model takes into account the nonlinear hysteretic behaviour of the structure through lumped plasticity beam-column elements that are modelled by using the modified Ibarra-Medina-Krawinkler nonlinear hinge model (Ibarra, Medina et al. 2005; Lignos and Krawinkler 2010) with degrading strength and stiffness characteristics. The monotonic moment-rotation relationship of this model is shown in Figure 4. The backbone curve and the hysteretic rules of this plastic hinge model are determined by six key parameters: the effective initial stiffness (K_e), the yield strength (M_y), the post-yield strength ratio (M_c /M_y), the plastic rotation capacity (θ_{pl}), the post capping rotation (θ_{pc}) and the cyclic deterioration (λ). These parameters are assumed to be Beta random variables with mean and coefficient of variation summarized in Table 1. The lower and upper bounds of these uncertain parameters were chosen broad enough such that they cover their expectable range (e.g. K_e varies between $0.15*K_g$ and $0.80*K_g$, for definition of K_g see Table 1).

The four-bay planar frame under consideration consists of ten different, six column and four beam, cross sections (i.e. potential plastic hinge locations) with different reinforcement ratios and/or axial loads. Hence, if each of the six random variables in Table 1 would have been assessed as independent, then a total number of sixty uncertain parameters associated with the lumped plasticity beam-column elements should have been considered. However, the latter would raise significant computational obstacles to the development of an efficient and accurate kriging surrogate model, since it would significantly increase the number of the predictor parameters in **x** described in the previous

section. Therefore, a further reduction of the number of uncertain parameters under consideration is achieved by assuming perfect correlations between variables associated with *strength*/*stiffness* (i.e. K_e, M_y, M_c /M_y) and *ductility* (i.e. θ_{pl}, θ_{pc}, λ) for each one of the ten different potential plastic hinges. Thus, under this assumption only twenty independent variables need to be considered, namely: *column strength*/*stiffness* ($c_{s,nc}$), *column ductility* ($c_{d,nc}$), *beam strength*/*stiffness* ($b_{s,nb}$) and *beam ductility* ($b_{d,nb}$), where nc = 1,…,6 and nb = 1,…,4 are indices denoting the different column and beam plastic hinges, respectively. Note that a similar approach for the reduction of the uncertain parameters associated with the lumped plasticity beam-column elements was used in (Liel, Haselton et al. 2009). Finally, the structural model is assumed to have Rayleigh damping with damping ratio ζ associated with the first and third mode shapes. The damping ratio ζ is assumed to be a lognormal variable with median value 5% and coefficient of variation equal to 30%. The fundamental period of the structure calculated based on the nominal (mean) values for K_e is equal to 1.13 s.

Table 1. Uncertainties in modeling parameters for lumped plasticity beam-column elements

Random variable	Mean	Coefficient of variation
Effective initial stiffness (K_e)	0.345*K_g*	0.280
Yield strength (M_y)	1.0*(computed+)	0.150
Post-yield strength ratio (M_c /M_y)	1.17	0.038
Plastic rotation capacity (θ_{pl})	1.0*(computed§)	0.410
Post capping rotation (θ_{pc})	0.10	0.410
Cyclic deterioration (λ)	1.0*(computed§)	0.370

* K_g: uncracked gross section of RC beam or column
+ Computed using moment-curvature fiber analysis with expected values for material properties
§ Computed using calibrations to experimental tests (Haselton and Deierlein 2007)

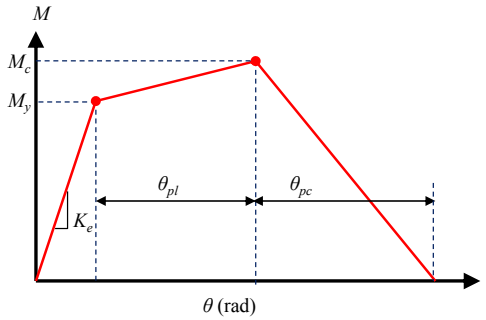

Figure 4. Monotonic moment-rotation for the modified Ibarra-Medina-Krawinkler deterioration model

The uncertain model parameters for the excitation include the moment magnitude and rupture distance as well as the characteristics for the near-fault pulse. Seismic events within range [M_{min}, M_{max}] = [5.0, 8.0] are considered for M, with uncertainty ultimately described (for the risk assessment) by the Gutenberg-Richter relationship leading to the PDF $p(M)=b_M\exp(-b_M M)/[\exp(-b_M M_{min})-\exp(-b_M M_{max})]$, where parameter b_M is selected to be $b_M=0.9\log_e(10)$. For the amplitude and period of the pulse the relationships (4) and (5) are adopted whereas the probability models for γ_p and ν_p are chosen as Gaussian with mean 1.8 and standard deviation 0.3 and uniform on the range [0, π], respectively. Regarding the uncertainty in the event location, the rupture distance, r, for the earthquake events is assumed to follow a log-normal distribution with median values r_{med} = 15 km and coefficient of variation 40%.

Ultimately the group of uncertain model parameters is $\boldsymbol{\theta}$=[M, r, A_p, T_p, γ_p, ν_p, $c_{s,nc}$, $c_{d,nc}$, $b_{s,nb}$, $b_{d,nb}$, ζ] when the near-fault pulse effect is incorporated in the ground motion model and $\boldsymbol{\theta}$=[M, r, $c_{s,nc}$, $c_{d,nc}$, $b_{s,nb}$, $b_{d,nb}$, ζ] when only the broadband far-field part of the excitation is considered, and a separate kriging model is developed for each of these two distinct cases as discussed previously. Space filling latin hypercube is selected for the support points in the range that are expected to take values based on the assumed probability models. For the bounded variables this corresponds to the truncation range whereas for the un-bounded variables this is defined to extend four standard deviations away from

their median values. A total of 9,000 points is used for P and 2,000 for NP. The larger number for P is motivated by challenges encountered when the directivity pulses induce resonance conditions (see also discussion later). The response quantities z_j predicted from the metamodel correspond to the peak interstory drifts (j=1,…,4) and peak floor accelerations (j=5,…,8, e.g. peak floor acceleration of the third floor is denoted as z_7) for the four stories of the structure.

The accuracy of the developed surrogate model is evaluated by calculating different error statistics using a leave-one-out cross-validation approach. This approach is established by removing sequentially each of the observations from the database, using the remaining support points to predict the output for that one and then evaluating the error between the predicted and real responses. The validation statistics are then obtained by averaging the errors established over all observations. The accuracy established is ultimately high with coefficient of determination over 94% and 99% for most approximated response quantities obtained from the metamodels developed for P and NP, respectively. These high values for the coefficient of determination indicate that the kriging model can describe very well the variability within the initial database, showing that it should be deemed appropriate for providing accurate estimates within a risk assessment framework.

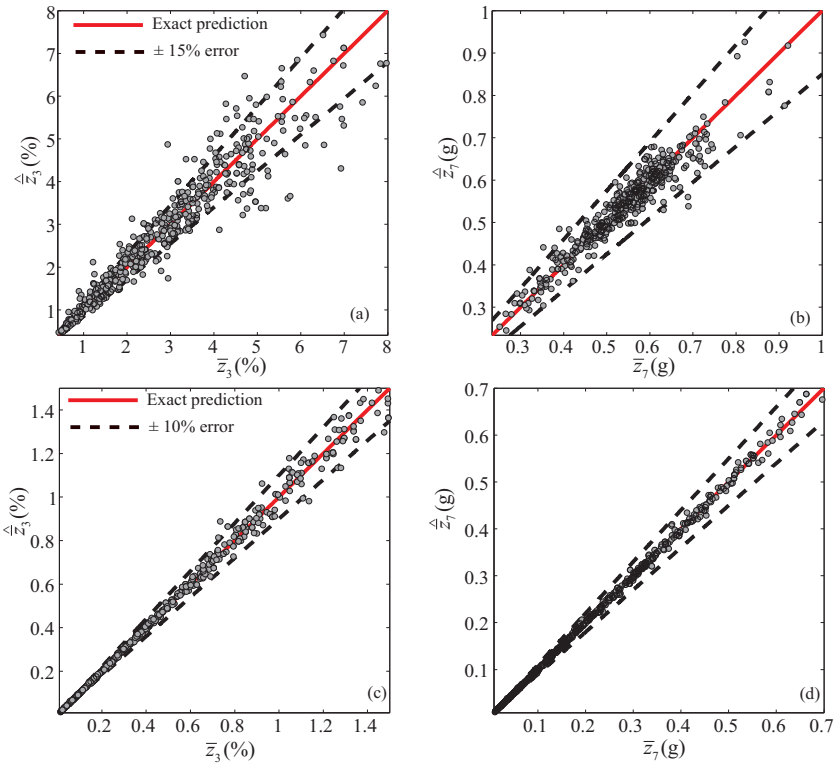

Figure 5. Comparison between high-fidelity model (OpenSees) results and kriging metamodel for the median of the peak interstory drift [parts (a), (c)] and the peak floor acceleration [parts (b), (d)] of the third floor. Top [parts (a), (b)] and bottom row [parts (c), (d)] correspond to P and NP models, respectively.

The prediction accuracy for the kriging model is also demonstrated in Figure 5, which shows the comparison between the median of the interstory drift of the third floor $\overline{z}_3$ [parts (a), (c)] and the corresponding median of the floor acceleration $\overline{z}_7$ [parts (b), (d)] obtained from OpenSees (high-fidelity model) and their kriging approximations $\hat{\overline{z}}_3$, $\hat{\overline{z}}_7$ for the P [parts (a), (b)] and NP [parts (c), (d)] seismic hazard models considered. It can be observed that in general the accuracy achieved through the kriging predictor is high. In particular, the accuracy of the acceleration medians appears to be higher than the one obtained for the drifts, a result that is attributed to the nonlinear hysteretic structural behavior that results in larger variability for the drift response, and as such greater challenges for any surrogate modeling approach to approximate that variability.

Regarding, now, the different accuracy levels achieved between the surrogate models for P and NP, it is observed that the kriging predictor developed for NP attains significantly higher accuracy comparing to the one that corresponds to P. This trend should be attributed to the fact that inclusion of the near fault pulse results in resonance conditions when the pulse period T_p lies in the vicinity of the fundamental structural period. In this case the structure exhibits severe inelastic response and consequently creates significantly large variability in the response quantities obtained from relatively similar vectors of input parameters **x**. This variability ultimately makes the development of a metamodel for P with similar accuracy as NP challenging. In this case if further improvement of the kriging model accuracy is sought, then the previous challenge can be addressed either by developing a kriging model based on a denser grid of support points or alternatively by implementing adaptive design of experiments techniques (Picheny, Ginsbourger et al. 2010).

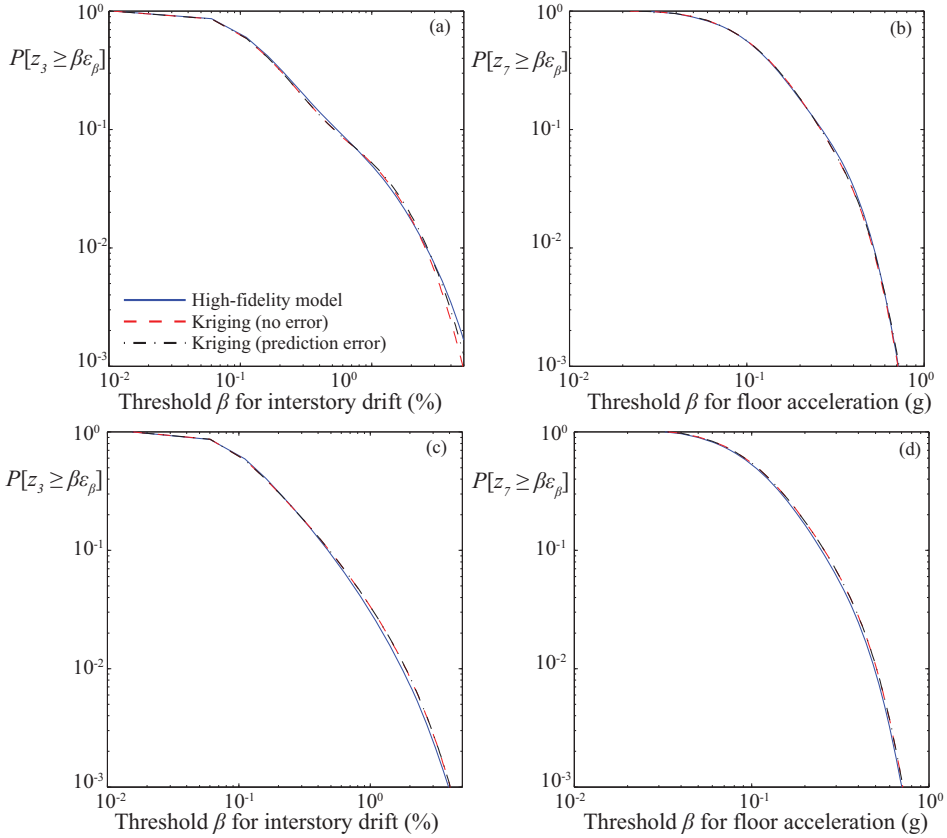

Figure 6. Probability $P(z_j \geq \beta\varepsilon_\beta)$ of exceeding a specific threshold based on high-fidelity model (OpenSees) and kriging metamodel for peak interstory drift [parts (a), (c)] and the peak floor acceleration [parts (b), (d)] of the third floor. Top [parts (a), (b)] and bottom row [parts (c), (d)] correspond to P and NP models, respectively.

Next, the optimized kriging metamodel is utilized for estimating seismic risk through (14) using N = 10,000 samples. This estimate is compared against the corresponding risk calculated through the high-fidelity model in OpenSees through (1). Figure 6 shows the probability of exceeding a specific threshold for the interstory drift [parts (a), (c)] and for floor acceleration [parts (b), (d)] of the third story for P [parts (a), (b)] and NP ([parts (c), (d)] hazard descriptions considered. The results correspond to σ_β = 0.2, whereas the average coefficient of variation (from the stochastic simulation) for all cases is not higher than 3.57%, corresponding to high-accuracy estimates (and thus appropriate for the comparison discussions here). Kriging predictions corresponding to inclusion or not of the model prediction error are reported. For the former risk is calculated using the formulation for the risk consequence measure $h(\boldsymbol{\theta})$ in (7), whereas for the latter formulation (13) is used. The risk corresponding to the high-fidelity model is estimated using $h(\boldsymbol{\theta})$ as shown in (2). The comparison

between the probabilities obtained from the high-fidelity model and the kriging metamodel indicates that the accuracy achieved is very high, leading to very small differences for all the cases. Although, as it was expected [see also Figure 5(a)] the bigger discrepancies between the OpenSees model and the kriging approximation correspond to the results related with drift response and the seismic hazard model that includes the near-fault effects, these deviations are considered to be relatively small. Furthermore, it can be observed that the inclusion of the kriging model prediction error only affects the surrogate model approximation of the risk related with the drift response of the P hazard model [Figure 6(a)], whereas the impact for the rest of the cases is negligible. In particular, it is observed that incorporation of the prediction error improves the metamodel based risk approximation that corresponds to small probabilities of exceedance (i.e. rare/high consequence events). This reveals the importance of incorporating the model prediction error when approximating risk related to rare events, while at the same time the negligible impact on improving the surrogate model approximation of more frequent events indicates that good accuracy risk estimates could have been obtained with smaller number of support points for these cases.

The total CPU time for this seismic risk assessment is only 265 s (P case) [79 s for NP case]. This is a huge reduction of computational time compared to the high fidelity model which requires 36.75 h (P case) [34.3 h for NP case]. Therefore, the surrogate modeling implementation presented in this study can support a highly efficient and accurate seismic risk assessment, and at the same time it allows the calculation of very accurate stochastic simulation estimates of seismic risk through (14), since a large number of samples N can be used with negligible computational effort.

Moving now to the investigation of the inclusion of the near-fault pulse effects on seismic risk, Figure 7 presents the comparison of the probability $P(z_j \geq \beta\varepsilon_\beta)$ of exceeding a specific threshold based on the kriging metamodel for peak interstory drift of the first floor [part (a)] and the peak floor acceleration [part (b)] of the fourth floor corresponding to P and NP seismic hazard descriptions. It is apparent that incorporation of the directivity pulses in the ground motion model drastically increases the probability of unacceptable performance related to drift response beyond a threshold of $\beta = 0.2\%$. The latter result indicates that near-fault pulse effects need to be carefully considered when describing the seismic hazard. On the contrary, seismic risk associated with acceleration response does not appear to be sensitive regarding the incorporation or not of near-fault pulses in the excitation model. This trend is attributed to the fact that the inelastic response induced by the pulses affects mainly the peak displacement response.

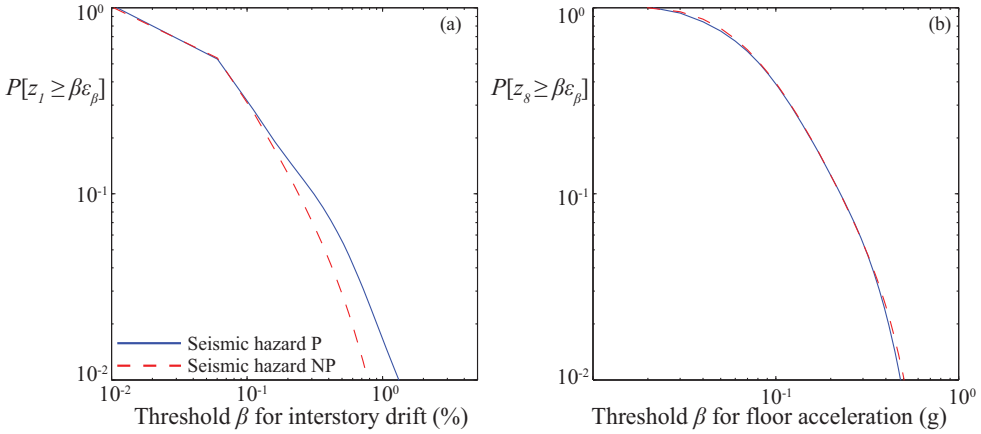

Figure 7. Comparison of the probability $P(z_j \geq \beta\varepsilon_\beta)$ of exceeding a specific threshold based on the kriging metamodel for peak interstory drift of the first floor [part (a)] and the peak floor acceleration [part (b)] of the fourth floor corresponding to P and NP seismic hazard descriptions.

CONCLUSIONS

An efficient computational framework, relying on kriging surrogate modeling, for estimating seismic risk with a modelling approach that utilizes stochastic ground motion models for the seismic hazard

characterization was presented in this paper. The surrogate model is used for approximating the structural response when addressing the uncertainty in the various model parameters and ultimately estimating seismic risk whereas the stochastic characteristics of the seismic excitation are addressed by assuming that under the influence of white noise each response quantity follows a lognormal distribution. The proposed framework is demonstrated for risk assessment of a benchmark reinforced concrete office building when the risk was characterized in terms of the fragility of different engineering demand parameters. Overall the study demonstrated the utility of the proposed kriging metamodel in supporting a highly efficient and accurate estimation of seismic risk.

REFERENCES

Atkinson, G. M. and W. Silva (2000). "Stochastic modeling of California ground motions." *Bulletin of the Seismological Society of America* 90(2): 255-274.

Ditlevsen, O. and H. O. Madsen (1996). Structural reliability methods. New York, NY, J. Wiley & Sons.

Ellingwood, B. R. (2001). "Earthquake risk assessment of building structures." *Reliability Engineering & System Safety* 74(3): 251-262.

Goulet, C. A., C. B. Haselton, et al. (2007). "Evaluation of the seismic performance of code-conforming reinforced-concrete frame building-From seismic hazard to collapse safety and economic losses." *Earthquake Engineering and Structural Dynamics* 36(13): 1973-1997.

Halldórsson, B., G. P. Mavroeidis, et al. (2010). "Near-Fault and Far-Field Strong Ground-Motion Simulation for Earthquake Engineering Applications Using the Specific Barrier Model." *Journal of Structural Engineering* 137(3): 433-444.

Haselton, C. and G. Deierlein (2007). Assessing Seismic Collapse Safety of Modern Reinforced Concrete Frame Buildings. *PEER Report 2007/08, Pacific Engineering Research Center*, University of California, Berkeley, California.

Ibarra, L. F., R. A. Medina, et al. (2005). "Hysteretic models that incorporate strength and stiffness deterioration." *Earthquake Engineering & Structural Dynamics* 34(12): 1489-1511.

Jalayer, F. and C. Cornell (2009). "Alternative non-linear demand estimation methods for probability-based seismic assessments." *Earthquake Engineering & Structural Dynamics* 38(8): 951-972.

Liel, A. B., C. B. Haselton, et al. (2009). "Incorporating modeling uncertainties in the assessment of seismic collapse risk of buildings." *Structural Safety* 31(2): 197-211.

Lignos, D. G. and H. Krawinkler (2010). "Deterioration modeling of steel components in support of collapse prediction of steel moment frames under earthquake loading." *Journal of Structural Engineering* 137(11): 1291-1302.

Lophaven, S. N., Nielsen, H.B., and Sondergaard, J. (2002). DACE-A MATLAB Kriging Toolbox, Technical University of Denmark.

Mavroeidis, G. P., G. Dong, et al. (2004). "Near-fault ground motions, and the response of elastic and inelastic single-degree-of-freedom(SDOF)systems." *Earthquake Engineering and Structural Dynamics* 33(9): 1023-1049.

Mavroeidis, G. P. and A. P. Papageorgiou (2003). "A mathematical representation of near-fault ground motions." *Bulletin of the Seismological Society of America* 93(3): 1099-1131.

McKenna, F. (2011). "OpenSees: A Framework for Earthquake Engineering Simulation." *Computing in Science & Engineering* 13(4): 58-66.

Picheny, V., D. Ginsbourger, et al. (2010). "Adaptive designs of experiments for accurate approximation of a target region." *Journal of Mechanical Design* 132(7): 071008.

Porter, K. A., R. P. Kennedy, et al. (2006). Developing Fragility Functions for Building Components, Report to ATC-58. Redwood City, CA, Applied Technology Council.

Robert, C. P. and G. Casella (2004). Monte Carlo statistical methods. New York, NY, Springer.

Taflanidis, A. A. and J. L. Beck (2009). "Life-cycle cost optimal design of passive dissipative devices " *Structural Safety* 31(6): 508-522.

Taflanidis, A. A. and G. Jia (2011). "A simulation-based framework for risk assessment and probabilistic sensitivity analysis of base-isolated structures." *Earthquake Engineering & Structural Dynamics* 40(14): 1629-1651.

Taflanidis, A. A., C. Vetter, et al. (2013). "Impact of modeling and excitation uncertainties on operational and structural reliability of Tension Leg Platforms." *Applied Ocean Research* 43: 131–147.

Zhang, J. and R. O. Foschi (2004). "Performance-based design and seismic reliability analysis using designed experiments and neural networks." *Probabilistic Engineering Mechanics* 19(3): 259-267.