ETICS 2021: Propagation of epistemic uncertainties and global sensitivity analysis in seismic risk assessment

Clément Gauchy, Cyril Feau, Josselin Garnier
clement.gauchy@cea.fr; cyril.feau@cea.fr; josselin.garnier@polytechnique.edu
– September 16, 2021 –
Table of contents

Introduction and state of the art

Propagation of epistemic uncertainties and global sensitivity analysis

Applications and numerical results
Introduction and state of the art
Seismic probabilistic risk assessment (SPRA) is dedicated in estimating the safety of a mechanical structure subjected to seismic ground motions and consists in three main steps:\(^1\):

- Seismic hazard probability distribution on a given site: \( dh(a) = p(a) d\mu \)

- Seismic fragility curve estimation \( \Psi(a) \). By definition the conditional probability of failure of the structure given a seismic intensity of level \( A = a \).

- Our QoI: Final probability of failure:

\[
\Upsilon = \int \Psi(a) dh(a)
\]

---

\(^1\)Robert P. Kennedy. Risk based seismic design criteria. *Nuclear Engineering and Design*, 1999
Seismic fragility curve

\[ \Psi(a) = \mathbb{P}(Y > C | A = a) \]

- **\( Y \):** mechanical demand of the structure, obtained using time-consuming numerical simulations.

- **\( C \):** critical level where the structure is considered in a failure state.

- **\( A \):** *Intensity Measure* of a seismic ground motion. Scalar value representing the intensity of the temporal seismic signal.
In practice there is not enough seismic ground motions for fragility curve estimation.

---


3Sanaz Rezaeian and Armen Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 2010
In practice there is not enough seismic ground motions for fragility curve estimation.

Example: European Strong-Motion Database\textsuperscript{2}. 97 real seismic signals with magnitude $5.5 < M < 6.5$ and distance to the epicenter $0 < D < 20$ km.

\textsuperscript{2}NN Ambraseys, P Smit, R Berardi, D Rinaldis, F Cotton, and C Berge. Dissemination of european strongmotion data. cd-rom collection. european commission, directorate-general xii, environmental and climate programme. Technical report, ENV4-CT97-0397, Brussels, Belgium, 2000

\textsuperscript{3}Sanaz Rezaeian and Armen Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. \textit{Earthquake Engineering & Structural Dynamics}, 2010
In practice there is not enough seismic ground motions for fragility curve estimation.

Example: European Strong-Motion Database\(^2\). 97 real seismic signals with magnitude \(5.5 < M < 6.5\) and distance to the epicenter \(0 < D < 20\) km.

Modelization by a filtered and modulated white noise stochastic process (good “space filling” properties).\(^3\)

---


\(^3\)Sanaz Rezaeian and Armen Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 2010
The choice of the seismic intensity

- The seismic intensity is a scalar quantity derived from the temporal seismic accelerogram.

- There are criteria in Earthquake Engineering to choose a “good” seismic intensity:
  
  - **Efficiency**: $Y|A \approx Y|\text{Seismic hazard}$.
  
  - **Sufficiency**: $Y|A$ has a small variance.

- We consider the two most classical intensity measures: the peak ground acceleration (PGA) and the spectral acceleration.

  - **Spectral acceleration**: Maximal acceleration of a linear oscillator at a chosen natural frequency and damping ratio (often very high sufficiency).

---

The mechanical demand $Y$ is usually obtained from time-consuming numerical simulations.

This motivates the use of statistical metamodel\textsuperscript{5} or active learning strategy\textsuperscript{6}.

\textsuperscript{5}Bertrand Iooss and Loïc Le Gratiet. Uncertainty and sensitivity analysis of functional risk curves based on gaussian processes. 
Reliability Engineering & System Safety, 2019

\textsuperscript{6}Rémi Sainct, Cyril Feau, Jean-Marc Martinez, and Josselin Garnier. Efficient methodology for seismic fragility curves estimation by active learning on support vector machines. 
Structural Safety, 2020
Main objectives of the thesis

- Developing **active learning strategies** to alleviate the computational burden of SPRA\(^7\).

- Modeling, propagating and assessing the impact of **epistemic uncertainties** on SPRA’s related quantities of interests (QoIs) such as fragility curves, probability of failures...

---

Propagation of epistemic uncertainties
Industrial motivations

Uncertainties are divided in two groups:

- **Aleatory**: Natural variability of a physical phenomenon.

- **Epistemic**: Comes from the Greek word $\epsilon\pi\sigma\tau\eta\mu\eta$ (knowledge). Uncertainties resulting from a lack of knowledge.

---

8 Armen Der Kiureghian and Ove Ditlevsen. *Aleatory or epistemic? does it matter?* Structural Safety, 2009
Uncertainties are divided in two groups:

- **Aleatory**: Natural variability of a physical phenomenon.

- **Epistemic**: Comes from the Greek word επιστηµη (knowledge). Uncertainties resulting from a lack of knowledge.

This division is purely subjective.\(^8\)

---

\(^8\)Armen Der Kiureghian and Ove Ditlevsen. *Aleatory or epistemic? does it matter?* Structural Safety, 2009
Uncertainties are divided in two groups:

- **Aleatory**: Natural variability of a physical phenomenon.

- **Epistemic**: Comes from the Greek word επιστηµη (knowledge). Uncertainties resulting from a lack of knowledge

This division is purely subjective.\(^8\).

For SPRA in the nuclear industry, epistemic uncertainties are mostly the mechanical parameters of the structure (natural frequency, damping ratio,...). Aleatory uncertainties come from the seismic ground motions’ stochasticity.

\(^8\)Armen Der Kiureghian and Ove Ditlevsen. *Aleatory or epistemic? does it matter?* Structural Safety, 2009
Our goal is to assess the effect of mechanical model parameters uncertainties (e.g. epistemic uncertainties) on the Quantity of Interest (QoI).

- For $X$ a random vector of parameters, $\Psi(a, X) = \mathbb{P}(Y > C|A = a, X)$

- The random function $\Psi(., X)$ could be seen as a functional QoI.

- Due to computational burden, we replace the costly mechanical simulation by a Gaussian Process surrogate of the mechanical response $Y$. 
Let $Z(x), \ x \in \mathbb{R}^d$ a random process.

A random process is a **Gaussian process** if for a tuple $(x_1, \ldots, x_n)$, the random vector $(Z(x_1), \ldots, Z(x_n))$ is Gaussian.

The probability distribution of a Gaussian process is completely defined by:

- Its mean function $\mu(x) = \mathbb{E}[Z(x)]$.
- Its covariance function $k(x, \tilde{x}) = \mathbb{E}[(Z(x) - \mu(x))(Z(\tilde{x}) - \mu(\tilde{x}))]$

we note $Z(x) \sim \mathcal{GP}(\mu(x), k(x, \tilde{x}))$
Thanks to Gaussian conditioning theorem we have that:

\[
(Z(x_*)|Z(x_1) = z_1, \ldots, Z(x_n) = z_n) \sim \mathcal{GP}(\mu_c(x_*), k_c(x_*, \tilde{x}))
\]

where:

\[
\mu_c(x) = \mu(x) + r(x)^T K^{-1}(z - \mu)
\]

\[
k_c(x, \tilde{x}) = k(x, \tilde{x}) - r(x)^T K^{-1} r(x)
\]

\[
\mu = (\mu(x_1), \ldots, \mu(x_n))^T \text{ the experiment design mean vector.}
\]

\[
z = (z_1, \ldots, z_n)^T \text{ the value of the model output.}
\]

\[
r(x) = (k(x, x_1), \ldots, k(x, x_n))^T \text{ covariance between a point } x \text{ and the experimental design.}
\]

\[
K = (k(x_i, x_j))_{1\leq i, j \leq n} \text{ covariance matrix of the experimental design.}
\]

We have an analytical expression about the probability distribution of \(Z(x_*)\) given data.
Gaussian Process regression for functional risk curves estimation have been previously studied.\(^9\)

We propose the following regression model:

\[
Y = \beta_0 + \beta_1 \log(A) + Z(X) + \varepsilon,
\]

where \(Z\) a centered Gaussian Process with a tensorized 5/2 Matèrn kernel on the log-epistemic variables \(\log(X)\), \(\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)\) and \(Y\) a mechanical response (e.g. peak displacement), submitted to a nonlinear transform (e.g. logarithm or Box Cox transformation).

\(\beta_0, \beta_1\), the Matèrn covariance hyperparameters and the nugget effect are calibrated by the posterior mode on a training dataset. The parameters are distributed by weakly informative priors.

Given a dataset \(\mathcal{D}_n = (A_i, X_i, Y_i)_{1 \leq i \leq n}\), the predictive distribution is

\[
(Y|\mathcal{D}_n) \sim \mathcal{GP}(\mu, \Sigma + \sigma_\varepsilon^2 Id)\]

with \(\mu\) and \(\Sigma\) respectively the conditional mean and covariance matrix of the Gaussian Process.

---

Reliability Engineering & System Safety, 2019
The GP surrogate model allows us to compute rapidly the fragility curves
\( \Psi(a, X) = \mathbb{P}(Y > C | A = a, X). \)

The fragility curve \( \Psi(a, X) \) can be approximated using the predictive distribution of the
GP by \( \tilde{\Psi}(a, X) \):

\[
\tilde{\Psi}(a, X) = \Phi \left( \frac{\mu(a, X) - C}{\sqrt{\Sigma(X, X) + \sigma_\varepsilon^2}} \right),
\]

It is also possible to propagate the GP surrogate uncertainty in the fragility curve, by
only integrating the Gaussian noise \( \varepsilon \):

\[
\tilde{\Psi}(a, X) = \Phi \left( \frac{\tilde{Y}_{GP}(a, X) - C}{\sigma_\varepsilon} \right).
\]

where \( \tilde{Y}_{GP}(A, X) \) is the Gaussian Process prediction at \((A, X)\) without the Gaussian noise \( \varepsilon \).
Propagation of the GP uncertainty on the mean fragility curve.

The mean fragility curve $\Psi(a) = \mathbb{P}_X(Y > C|A = a)$ is approximated using our Gaussian process $\Psi_{GP}(a) = \mathbb{P}_X(Y_{GP}(a, X) > C|A = a)$. Thus, $\Psi_{GP}(a)$ is tainted with the uncertainties coming from the Gaussian process regression.
The mean fragility curve $\Psi(a) = \mathbb{P}_X(Y > C|A = a)$ is approximated using our Gaussian process $\Psi_{GP}(a) = \mathbb{P}_X(Y_{GP}(a, X) > C|A = a)$. Thus, $\Psi_{GP}(a)$ is tainted with the uncertainties coming from the Gaussian process regression.

$\Psi(a)$ can be approximated by the mean value $\tilde{\Psi}(a)$ of $\Psi_{GP}(a)$:

$$
\tilde{\Psi}(a) = \mathbb{E}_X \left[ \Phi \left( \frac{\mu(a, X) - C}{\sqrt{\Sigma(X, X) + \sigma_\epsilon^2}} \right) \right]
$$
Propagation of the GP uncertainty on the mean fragility curve.

- The mean fragility curve $\Psi(a) = \mathbb{P}_X(Y > C|A = a)$ is approximated using our Gaussian process $\Psi_{GP}(a) = \mathbb{P}_X(Y_{GP}(a, X) > C|A = a)$. Thus, $\Psi_{GP}(a)$ is tainted with the uncertainties coming from the Gaussian process regression.

- $\Psi(a)$ can be approximated by the mean value $\bar{\Psi}(a)$ of $\Psi_{GP}(a)$:

  $$
  \bar{\Psi}(a) = \mathbb{E}_X \left[ \Phi \left( \frac{\mu(a, X) - C}{\sqrt{\Sigma(X, X) + \sigma^2_{\epsilon}}} \right) \right]
  $$

- GP uncertainty is propagated using Monte Carlo on a dataset $(X_k)_{1 \leq k \leq N}$ and $L$ Gaussian process samples $(y^{(l)}(a, X_k))_{1 \leq l \leq L}$:

  $$
  \Psi^{(l)}_{GP}(a) = \frac{1}{N} \sum_{k=1}^{N} \mathbf{1}_{(y^{(l)}(a, X_{k}) > C)}
  $$
Propagation of the GP uncertainty on the mean fragility curve.

- The mean fragility curve $\Psi(a) = \mathbb{P}_X(Y > C|A = a)$ is approximated using our Gaussian process $\Psi_{GP}(a) = \mathbb{P}_X(Y_{GP}(a, X) > C|A = a)$. Thus, $\Psi_{GP}(a)$ is tainted with the uncertainties coming from the Gaussian process regression.

- $\Psi(a)$ can be approximated by the mean value $\bar{\Psi}(a)$ of $\Psi_{GP}(a)$:

\[
\bar{\Psi}(a) = \mathbb{E}_X \left[ \Phi \left( \frac{\mu(a, X) - C}{\sqrt{\Sigma(X, X) + \sigma^2_\varepsilon}} \right) \right]
\]

- GP uncertainty is propagated using Monte Carlo on a dataset $(X_k)_{1 \leq k \leq N}$ and $L$ Gaussian process samples $(y^{(\ell)}(a, X_k))_{1 \leq \ell \leq L}$:

\[
\Psi_{GP}^{(\ell)}(a) = \frac{1}{N} \sum_{k=1}^{N} 1_{(y^{(\ell)}(a, X_k) > C)}
\]

- $\Psi_{GP}^{(\ell)}(a)$ is thus a Monte Carlo approximation of a realization of the random variable $\Psi_{GP}(a)$. 
Sobol indices are variance-based importance measures coming from the Hoeffding-Sobol decomposition of the variance.

Aggregated first order and total order Sobol indices for functional risk curves have been previously studied\(^\text{10}\), raising the following expressions:

\[
S_i = \frac{\mathbb{E}[\|\tilde{\Psi} - \Psi_{X^i}\|^2_{L^2}]}{\mathbb{E}[\|\tilde{\Psi} - \Psi_X\|^2_{L^2}]},
\]

\[
T_i = \frac{\mathbb{E}[\|\tilde{\Psi} - \Psi_{X^{-i}}\|^2_{L^2}]}{\mathbb{E}[\|\tilde{\Psi} - \Psi_X\|^2_{L^2}]},
\]

with \(\tilde{\Psi} = \mathbb{E}[\Psi(\cdot, X)]\), \(\Psi_{X^i} = \mathbb{E}[\Psi(\cdot, X)|X^i]\) and \(X^{-i} = (X^1, \ldots, X^{i-1}, X^{i+1}, \ldots, X^p)\).

---

Lemma: Denote $\tilde{X}^i$ the random vector such that $(\tilde{X}_j)_i = X^i$ if $i = j$ and $(\tilde{X}_j)_i = \tilde{X}^i$ if $i \neq k$ where $\tilde{X}^i$ is an independent copy of $X^i$. Thus:

$$S_i = \frac{\mathbb{E}[\langle \Psi_X, \Psi_{\tilde{X}^i} \rangle_{L^2}]}{\mathbb{E}[\| \Psi - \Psi_X \|_{L^2}^2]} ,$$

$$T_i = 1 - \frac{\mathbb{E}[\langle \Psi_X, \Psi_{\tilde{X}^{-i}} \rangle_{L^2}]}{\mathbb{E}[\| \Psi - \Psi_X \|_{L^2}^2]} .$$

Where $\langle \Psi_X, \Psi_{\tilde{X}^i} \rangle_{L^2} = \int (\Psi_X(a) - \bar{\Psi}(a))(\Psi_{\tilde{X}^i}(a) - \bar{\Psi}(a))da$.

We propose a Monte Carlo estimator of the aggregated Sobol indices using the lemma above.
The GP uncertainty on the aggregated Sobol indices can be propagated (previously done for the case of scalar QoI\textsuperscript{11}).

Principle: replace $\Psi(., X)$ by $\tilde{\Psi}(., X)$. Thus we have to simulate GP trajectories $(\tilde{Y}_{GP}(a, X_k))_{1 \leq k \leq N}$ for $N$ large, which is computationally challenging. One can rely on approximation methods such as spectral methods or Karhunen-Loeve decomposition.

\textsuperscript{11}Loic Le Gratiet, Claire Cannamela, and Bertrand Iooss. A bayesian approach for global sensitivity analysis of (multifidelity) computer codes. *SIAM/ASA Journal on Uncertainty Quantification*, 2014
Applications and numerical results
Numerical application: Nonlinear oscillator

Figure: Elasto-plastic mechanical oscillator with kinematic hardening. The linear parameters are the mass $m$, stiffness $k$, damping ratio $\xi$. The non linearity is controlled by the yield limit $\ell_Y$ and the post-yield stiffness $a$.

\[
\ddot{z}(t) + 2\beta \omega_L \dot{z}(t) + f_{NL}(t) = -s(t),
\]

Table: Epistemic uncertainties on the elasto-plastic oscillator

<table>
<thead>
<tr>
<th>parameter</th>
<th>distribution</th>
<th>mean</th>
<th>c.o.v</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>Lognormal</td>
<td>1</td>
<td>30%</td>
</tr>
<tr>
<td>$k$</td>
<td>Lognormal</td>
<td>900</td>
<td>30%</td>
</tr>
<tr>
<td>$\xi$</td>
<td>Lognormal</td>
<td>0.02</td>
<td>50%</td>
</tr>
<tr>
<td>$\ell_Y$</td>
<td>Lognormal</td>
<td>$5 \times 10^{-3}$</td>
<td>30%</td>
</tr>
<tr>
<td>$a$</td>
<td>Lognormal</td>
<td>0.2</td>
<td>30%</td>
</tr>
</tbody>
</table>
The intensity measure chosen is the *Spectral acceleration* at 5Hz and 2\% damping ratio.

For a seismic signal $t \rightarrow s_i(t)$, the oscillator response $t \rightarrow z_i(t)$ is computed, then the failure state is defined by:

$$\max_{t \in [0,T]} |z_i(t)| > 2\ell_Y$$
The metamodel is trained on an optimized LHS with according probability distribution for each variable using maximin criterion (PyDOE package in Python) of 500 datapoints.

The input variables are standardized (centered and scaled).

Performance of the metamodel on the mechanical demand using the mean $Q^2$ predictivity coefficient computed with 5-fold cross validation on an optimized LHS dataset of size 500.

$$Q^2 = 0.86$$
90% and 10% quantiles fragility curves estimated on a Monte Carlo of size $N = 10000$ tainted with epistemic and GP uncertainties for the nonlinear oscillator. The metamodel is trained with 80 observations.
90% and 10% quantiles for the mean fragility curves estimated on a Monte Carlo of size $N = 5000$ and $L = 3000$ GP trajectories drawn. The metamodel is trained with an optimized LHS of 50 observations.
Aggregated Sobol indices estimation with posterior mean of a GP metamodel

First order Sobol indices estimation on $B = 20$ bootstrap samples of size $N = 15000$. GP metamodel is trained on an optimized LHS of 50 observations.
Total order Sobol indices estimation on $B = 20$ bootstrap samples of size $N = 15000$. GP metamodel is trained on an optimized LHS of 50 observations.
Aggregated Sobol indices estimation with a GP metamodel

First order Sobol indices estimation on $B = 30$ bootstraps samples of size $N = 15000$ and $M = 50$ draws of the GP prediction $\tilde{Y}_{GP}(A, X)$. GP metamodel is trained on a optimized LHS of 50 observations.
Industrial test case: Safety water pipe of a French PWR

- The following test case corresponds to a piping system which is a simplified part of a secondary line of a French Pressurized Water Reactor (ASG program).

- The CAST3M numerical model was validated based on seismic tests performed on the shaking table Azalee of the EMSI laboratory of CEA Saclay.

- The failure state is defined by the out-of-plane rotation of the elbow near the clamped end of the pipe. Failure is defined by a rotation that exceeds 1°.

Overview of the ASG mock-up on the shaking table
Uncertainties on the boundary conditions has been introduced to simulate the interaction between the simplified piping system and the overall structure. We also suppose the piping structure is inside a fictitious building with unknown mechanical properties (a natural frequency and a damping ratio).

The 12 uncertain parameters have uniform distribution with a coefficient of variation of 15%.
90% and 10% quantiles fragility curves estimated on a Monte Carlo of size $N = 10000$ tainted with epistemic and GP uncertainties for the ASG piping structure simulation (linear analysis). The metamodel is trained with 200 observations.
Conclusion & perspectives

- Uncertainty propagation framework for seismic risk assessment using Gaussian process regression, taking into account the metamodel uncertainties.

- Global Sensitivity Analysis framework on fragility curves.

**What’s next ?:**
- Multifidelity metamodel for the ASG piping system.
- Defining insightful QoIs for SPRA.
- Kernel based GSA (e.g. MMD sensitivity indices).
Thank you for your attention!
Références


Merci pour votre attention !