PhD Defense

Uncertainty quantification methodology for seismic fragility curves of mechanical structures - Application to a piping system of a nuclear power plant

Clément GAUCHY (CEA, CMAP), Cyril FEAU (CEA), Josselin GARNIER (CMAP)

Seismic probabilistic risk assessment

Seismic probabilistic risk assessment (SPRA) is dedicated to estimating the safety of a mechanical structure subjected to seismic ground motions and consists in three main steps [1]:

- **1)** Estimation of the probability measure $\boldsymbol{\lambda}$ of annual occurence of a seismic ground motion of intensity $\boldsymbol{a} \in \mathbb{R}^+$. $\hookrightarrow$ Probabilistic Seismic Hazard Analysis (PSHA);

[1] Robert P. Kennedy. Risk based seismic design criteria. *Nuclear Engineering and Design*, 1999

Seismic probabilistic risk assessment

Seismic probabilistic risk assessment (SPRA) is dedicated to estimating the safety of a mechanical structure subjected to seismic ground motions and consists in three main steps [1]:

- **1)** Estimation of the probability measure $\boldsymbol{\lambda}$ of annual occurence of a seismic ground motion of intensity $\boldsymbol{a} \in \mathbb{R}^+$. $\hookrightarrow$ Probabilistic Seismic Hazard Analysis (PSHA);
- **2)** The estimation of the probability of failure $\psi(\boldsymbol{a})$ of a structure conditional to the seismic intensity $\boldsymbol{a}$. ψ is the **seismic fragility curve**;

[1] Robert P. Kennedy. Risk based seismic design criteria. *Nuclear Engineering and Design*, 1999

Seismic probabilistic risk assessment

Seismic probabilistic risk assessment (SPRA) is dedicated to estimating the safety of a mechanical structure subjected to seismic ground motions and consists in three main steps [1]:

- **1)** Estimation of the probability measure $\boldsymbol{\lambda}$ of annual occurence of a seismic ground motion of intensity $\boldsymbol{a} \in \mathbb{R}^+$. $\hookrightarrow$ Probabilistic Seismic Hazard Analysis (PSHA);
- **2)** The estimation of the probability of failure $\psi(\boldsymbol{a})$ of a structure conditional to the seismic intensity $\boldsymbol{a}$. ψ is the **seismic fragility curve**;
- **3)** The estimation of the annual probability of failure of the structure:

$$\boldsymbol{P_f} = \int \psi(\boldsymbol{a}) d\boldsymbol{\lambda}(\boldsymbol{a}) \,.$$

[1] Robert P. Kennedy. Risk based seismic design criteria. *Nuclear Engineering and Design*, 1999

Seismic probabilistic risk assessment

Seismic probabilistic risk assessment (SPRA) is dedicated to estimating the safety of a mechanical structure subjected to seismic ground motions and consists in three main steps [1]:

- **1)** Estimation of the probability measure $\boldsymbol{\lambda}$ of annual occurence of a seismic ground motion of intensity $\boldsymbol{a} \in \mathbb{R}^+$. $\hookrightarrow$ Probabilistic Seismic Hazard Analysis (PSHA);
- **2)** The estimation of the probability of failure $\psi(\boldsymbol{a})$ of a structure conditional to the seismic intensity $\boldsymbol{a}$. ψ is the **seismic fragility curve**;
- **3)** The estimation of the annual probability of failure of the structure:

$$\boldsymbol{P_f} = \int \psi(\boldsymbol{a}) d\boldsymbol{\lambda}(\boldsymbol{a}) \,.$$

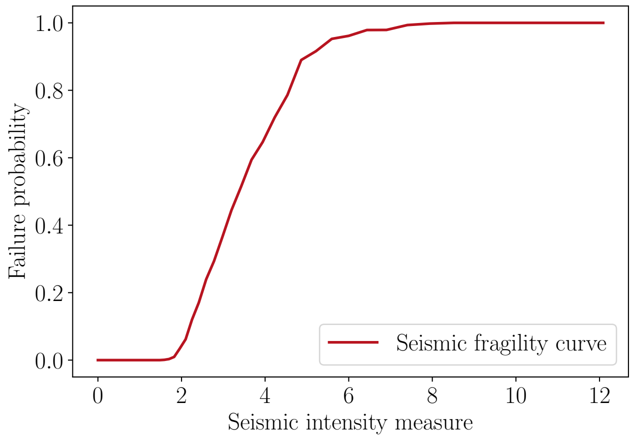
$\hookrightarrow$ This thesis focuses on the second step.

[1] Robert P. Kennedy. Risk based seismic design criteria. *Nuclear Engineering and Design*, 1999

Seismic fragility curve

$$\psi(\boldsymbol{a}) = \mathbb{P}(\boldsymbol{Z} > \boldsymbol{C} | \boldsymbol{A} = \boldsymbol{a})$$

- $\boldsymbol{Z}$: Mechanical demand of the structure, obtained using time-consuming numerical simulations;
- $\boldsymbol{C}$: Critical level for which the structure is considered in a failure state;
- $\boldsymbol{A}$: Intensity measure of a seismic ground motion. Scalar value representing the intensity of the temporal seismic signal.

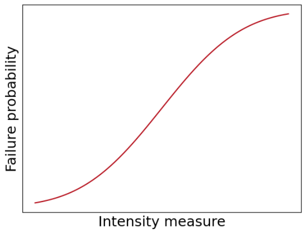


Classical estimation framework

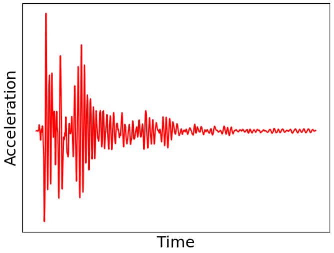
Mechanical parameters of the structure
$\mathbf{x} = (E, \omega, \xi, \dots)$

Computer model of a structure

Seismic fragility curve



Seismic ground motions



Expert judgement estimation

The first approach proposed in the 1980s for seismic fragility curves evaluation is based on expert judgements [2].

The seismic capacity $\boldsymbol{A_c}$ is defined by

$$\boldsymbol{A_c} = \boldsymbol{A_m} \boldsymbol{\varepsilon_R} \boldsymbol{\varepsilon_U} \ ,$$

where $\boldsymbol{A_m}$ is the median capacity, $\boldsymbol{\varepsilon_R}$ and $\boldsymbol{\varepsilon_U}$ follow lognormal distributions with unit median.

[2] R.P. Kennedy, C.A. Cornell, R.D. Campbell, S. Kaplan, and H.F. Perla. Probabilistic seismic safety study of an existing nuclear power plant.
Nuclear Engineering and Design, 59(2):315 – 338, 1980

Expert judgement estimation

The first approach proposed in the 1980s for seismic fragility curves evaluation is based on expert judgements [2].

The seismic capacity $\boldsymbol{A_c}$ is defined by

$$\boldsymbol{A_c = A_m \varepsilon_R \varepsilon_U} \ ,$$

where $\boldsymbol{A_m}$ is the median capacity, $\boldsymbol{\varepsilon_R}$ and $\boldsymbol{\varepsilon_U}$ follow lognormal distributions with unit median.

$\boldsymbol{\varepsilon_R}$ $\rightarrow$ Uncertainty from the seismic ground motions.
$\boldsymbol{\varepsilon_U}$ $\rightarrow$ Uncertainty coming from a lack of knowledge on the system.

[2] R.P. Kennedy, C.A. Cornell, R.D. Campbell, S. Kaplan, and H.F. Perla. Probabilistic seismic safety study of an existing nuclear power plant.
Nuclear Engineering and Design, 59(2):315 – 338, 1980

Expert judgement estimation

The first approach proposed in the 1980s for seismic fragility curves evaluation is based on expert judgements [2].

The seismic capacity $\boldsymbol{A_c}$ is defined by

$$\boldsymbol{A_c} = \boldsymbol{A_m}\boldsymbol{\varepsilon_R}\boldsymbol{\varepsilon_U} \ ,$$

where $\boldsymbol{A_m}$ is the median capacity, $\boldsymbol{\varepsilon_R}$ and $\boldsymbol{\varepsilon_U}$ follow lognormal distributions with unit median.

$\boldsymbol{\varepsilon_R}$ $\rightarrow$ Uncertainty from the seismic ground motions.
$\boldsymbol{\varepsilon_U}$ $\rightarrow$ Uncertainty coming from a lack of knowledge on the system.

Fragility curve evaluation

$$\psi(\boldsymbol{a}) = \mathbb{P}(\boldsymbol{A_c} \leq \boldsymbol{a}) \ .$$

[2] R.P. Kennedy, C.A. Cornell, R.D. Campbell, S. Kaplan, and H.F. Perla. Probabilistic seismic safety study of an existing nuclear power plant.
Nuclear Engineering and Design, 59(2):315 – 338, 1980

Motivation of the thesis

The sources of uncertainties are divided in two groups[3].

[3] A. Der Kiureghian and O. Ditlevsen. Aleatory or epistemic? does it matter? *Structural Safety*, 31(2):105–112, 2009

Motivation of the thesis

The sources of uncertainties are divided in two groups[3].

- The **aleatory uncertainties** coming from the natural variability of physical phenomena;

[3] A. Der Kiureghian and O. Ditlevsen. Aleatory or epistemic? does it matter? *Structural Safety*, 31(2):105–112, 2009

Motivation of the thesis

The sources of uncertainties are divided in two groups[3].

- The **aleatory uncertainties** coming from the natural variability of physical phenomena;

- The **epistemic uncertainties** coming from a lack of knowledge of the system studied. **They can be reduced in the short term (data gathering, experts knowledge,...);**

[3] A. Der Kiureghian and O. Ditlevsen. Aleatory or epistemic? does it matter? *Structural Safety*, 31(2):105–112, 2009

Motivation of the thesis

The sources of uncertainties are divided in two groups[3].

- The **aleatory uncertainties** coming from the natural variability of physical phenomena;
- The **epistemic uncertainties** coming from a lack of knowledge of the system studied. **They can be reduced in the short term (data gathering, experts knowledge,...);**
- This is a purely practical separation of uncertainties sources that is useful for engineering studies;

[3] A. Der Kiureghian and O. Ditlevsen. Aleatory or epistemic? does it matter? *Structural Safety*, 31(2):105–112, 2009

Motivation of the thesis

The sources of uncertainties are divided in two groups[3].

- The **aleatory uncertainties** coming from the natural variability of physical phenomena;
- The **epistemic uncertainties** coming from a lack of knowledge of the system studied. **They can be reduced in the short term (data gathering, experts knowledge,...);**
- This is a purely practical separation of uncertainties sources that is useful for engineering studies;
- The effects of epistemic uncertainties in fragility analysis are usually assessed using expert judgements.

[3] A. Der Kiureghian and O. Ditlevsen. Aleatory or epistemic? does it matter? *Structural Safety*, 31(2):105–112, 2009

Motivation of the thesis

The sources of uncertainties are divided in two groups[3].

- The **aleatory uncertainties** coming from the natural variability of physical phenomena;
- The **epistemic uncertainties** coming from a lack of knowledge of the system studied. **They can be reduced in the short term (data gathering, experts knowledge,...);**
- This is a purely practical separation of uncertainties sources that is useful for engineering studies;
- The effects of epistemic uncertainties in fragility analysis are usually assessed using expert judgements.

Goal: Data-driven assessment of epistemic uncertainties on seismic fragility curves using computer simulations → Uncertainty Quantification (UQ) methodology.

[3] A. Der Kiureghian and O. Ditlevsen. Aleatory or epistemic? does it matter? *Structural Safety*, 31(2):105–112, 2009

Outline of the defense

1. Uncertainty Quantification
2. Surrogate modeling of computer codes using Gaussian process
3. Global sensitivity analysis on seismic fragility curves
4. Sequential design of experiments
5. Conclusion and perspectives

Outline of the defense

1. Uncertainty Quantification
   - General framework
   - Extension to earthquake engineering
2. Surrogate modeling of computer codes using Gaussian process
3. Global sensitivity analysis on seismic fragility curves
4. Sequential design of experiments
5. Conclusion and perspectives

UQ is an interdisciplinary framework aiming to assess the uncertainties tainting a complex engineering system [4]. It has several objectives:

[4] E. De Rocquigny, N. Devictor, S. Tarantola, Y. Lefebvre, N. Pérot, W. Castaings, F. Mangeant, C. Schwob, R. Lavin, J.R. Masse, P. Limbourg, W. Kanning, and P.H.A.J.M. Gelder. *Uncertainty in industrial practice: A guide to quantitative uncertainty management.* Wiley, 2008

UQ is an interdisciplinary framework aiming to assess the uncertainties tainting a complex engineering system [4]. It has several objectives:

- Optimize the exploitation under costs and risks constraints;

[4] E. De Rocquigny, N. Devictor, S. Tarantola, Y. Lefebvre, N. Pérot, W. Castaings, F. Mangeant, C. Schwob, R. Lavin, J.R. Masse, P. Limbourg, W. Kanning, and P.H.A.J.M. Gelder. *Uncertainty in industrial practice: A guide to quantitative uncertainty management*. Wiley, 2008

UQ is an interdisciplinary framework aiming to assess the uncertainties tainting a complex engineering system [4]. It has several objectives:

- Optimize the exploitation under costs and risks constraints;
- Estimate a regulatory criterion;

[4] E. De Rocquigny, N. Devictor, S. Tarantola, Y. Lefebvre, N. Pérot, W. Castaings, F. Mangeant, C. Schwob, R. Lavin, J.R. Masse, P. Limbourg, W. Kanning, and P.H.A.J.M. Gelder. *Uncertainty in industrial practice: A guide to quantitative uncertainty management.* Wiley, 2008

UQ is an interdisciplinary framework aiming to assess the uncertainties tainting a complex engineering system [4]. It has several objectives:

- Optimize the exploitation under costs and risks constraints;
- Estimate a regulatory criterion;
- Acquire a better knowledge of the system under study.

[4] E. De Rocquigny, N. Devictor, S. Tarantola, Y. Lefebvre, N. Pérot, W. Castaings, F. Mangeant, C. Schwob, R. Lavin, J.R. Masse, P. Limbourg, W. Kanning, and P.H.A.J.M. Gelder. *Uncertainty in industrial practice: A guide to quantitative uncertainty management*. Wiley, 2008

Uncertainty Quantification

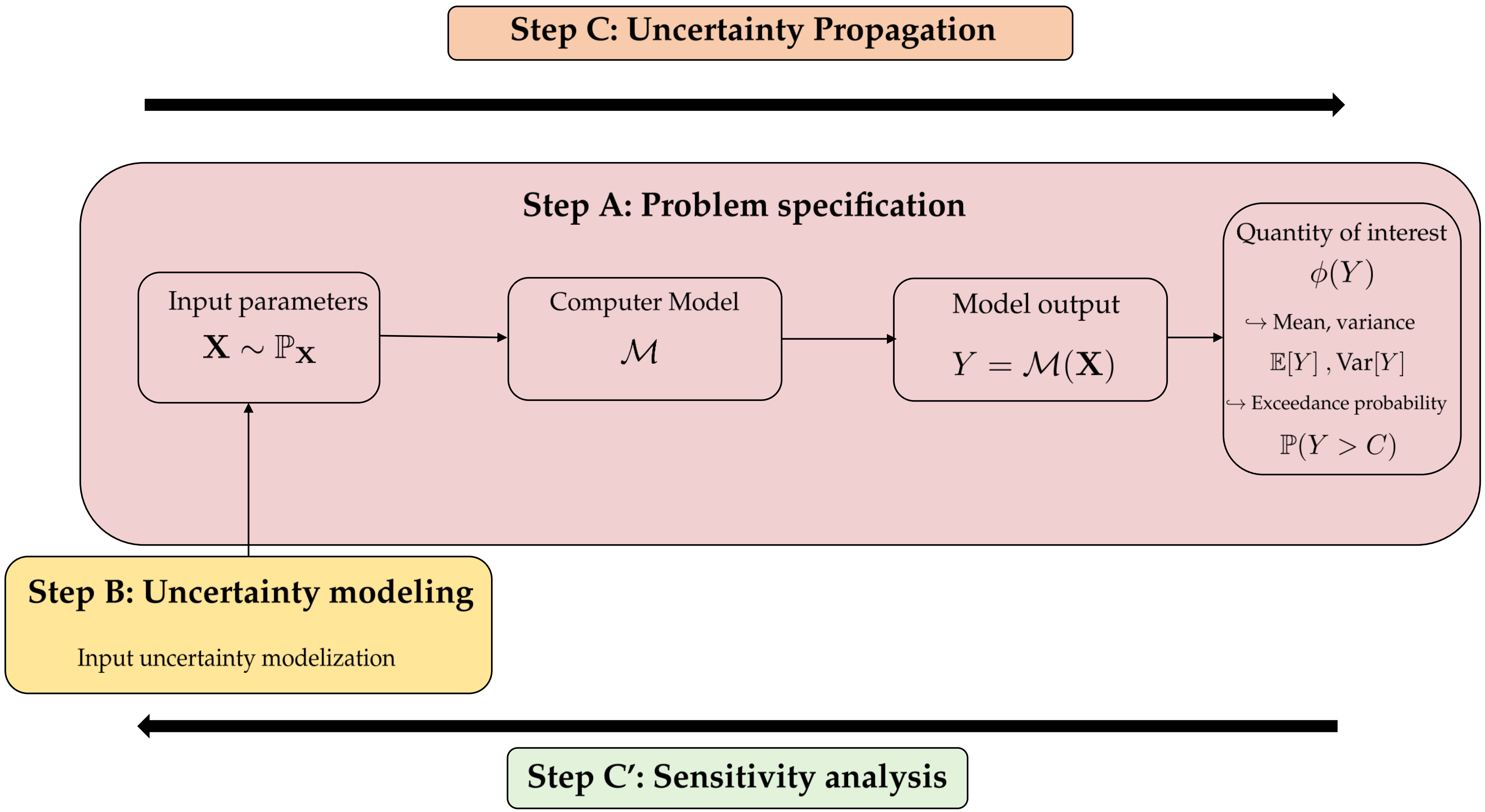
UQ is an interdisciplinary framework aiming to assess the uncertainties tainting a complex engineering system [4]. It has several objectives:

- Optimize the exploitation under costs and risks constraints;
- Estimate a regulatory criterion;
- Acquire a better knowledge of the system under study.

The computer model of the engineering system is modeled as a function $\mathcal{M} : \mathbb{R}^p \mapsto \mathbb{R}$

[4] E. De Rocquigny, N. Devictor, S. Tarantola, Y. Lefebvre, N. Pérot, W. Castaings, F. Mangeant, C. Schwob, R. Lavin, J.R. Masse, P. Limbourg, W. Kanning, and P.H.A.J.M. Gelder. *Uncertainty in industrial practice: A guide to quantitative uncertainty management*. Wiley, 2008

Uncertainty Quantification[5]



E. De Rocquigny, N. Devictor, S. Tarantola, Y. Lefebvre, N. Pérot, W. Castaings, F. Mangeant, C. Schwob, R. Lavin, J.R. Masse, P. Limbourg, W. Kanning, and P.H.A.J.M. Gelder. *Uncertainty in industrial practice: A guide to quantitative uncertainty management.* Wiley, 2008

Extension to earthquake engineering

- The motivation of the thesis is to perform Uncertainty Quantification on the seismic fragility curves;

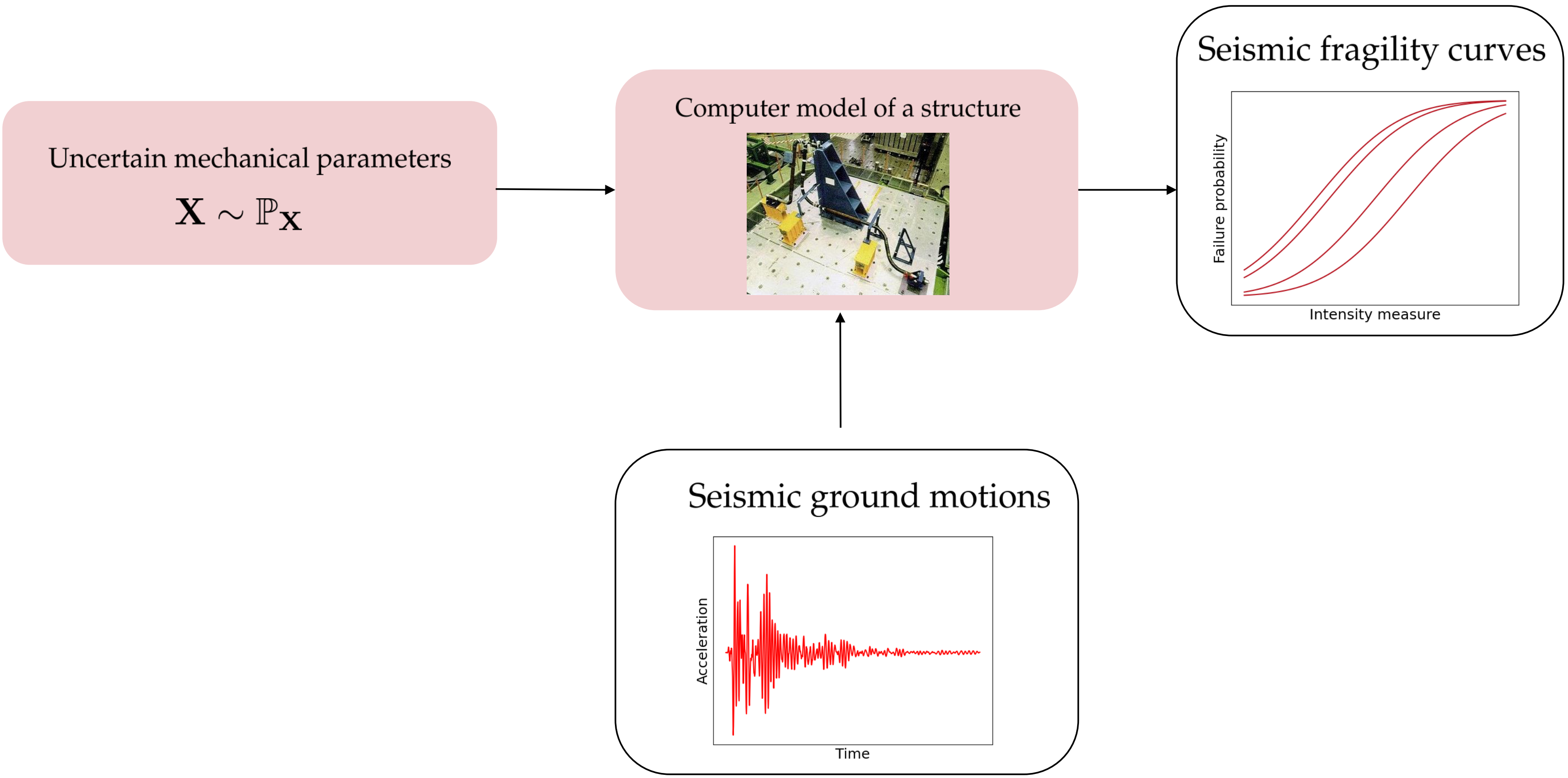
Extension to earthquake engineering

- The motivation of the thesis is to perform Uncertainty Quantification on the seismic fragility curves;
- We consider that the source of epistemic uncertainties comes from the uncertainties on the mechanical parameters of the structure;

Extension to earthquake engineering

- The motivation of the thesis is to perform Uncertainty Quantification on the seismic fragility curves;
- We consider that the source of epistemic uncertainties comes from the uncertainties on the mechanical parameters of the structure;
- $\hookrightarrow$ Our contribution: Adaptation of the UQ framework to earthquake engineering.

Proposed estimation framework



Seismic fragility curve with epistemic uncertainties

We propose a new definition of fragility curves encompassing mechanical parameters uncertainties.

$$\boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{X}) = \mathbb{P}\left(\boldsymbol{z}(\boldsymbol{A}, \mathrm{X}) > \boldsymbol{C} | \boldsymbol{A} = \boldsymbol{a}, \mathrm{X}\right)$$

$\boldsymbol{\Psi}$ is a random seismic fragility curve, whose randomness comes from the mechanical parameters uncertainties.

Seismic fragility curve with epistemic uncertainties

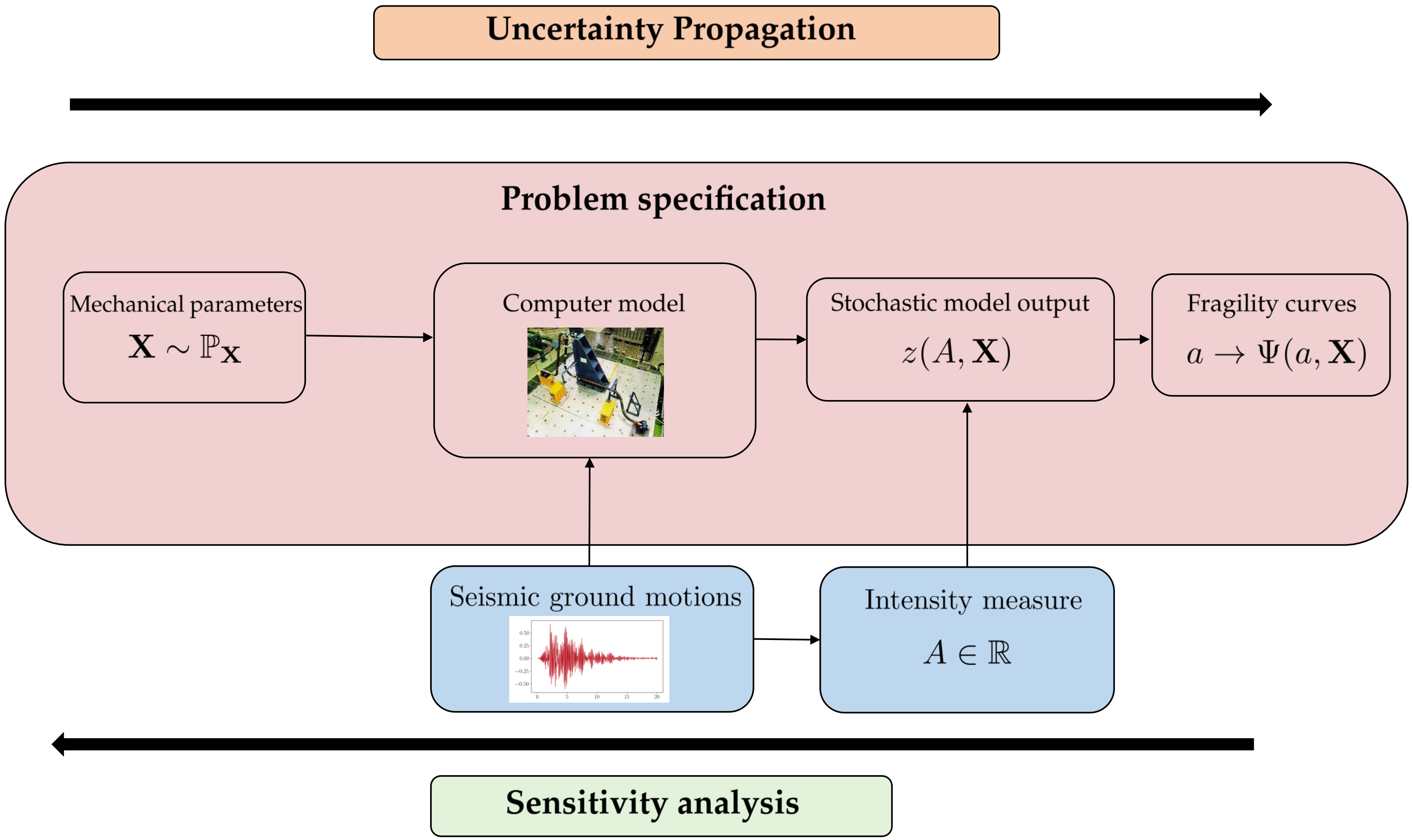
We propose a new definition of fragility curves encompassing mechanical parameters uncertainties.

$$\Psi(a, \mathrm{X}) = \mathbb{P}\left(z(A, \mathrm{X}) > C | A = a, \mathrm{X}\right)$$

Ψ is a random seismic fragility curve, whose randomness comes from the mechanical parameters uncertainties.

z is the stochastic model output. Its randomness comes from the seismic ground motion uncertainty.

Proposed UQ framework



Insightful quantities of interest

From the probability distribution of $\boldsymbol{a} \rightarrow \boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{X})$, we can derive statistical quantities of interest.

Insightful quantities of interest

From the probability distribution of $\boldsymbol{a} \to \boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{X})$, we can derive statistical quantities of interest.

- Classical quantity of interest: The mean seismic fragility curve

$$\bar{\boldsymbol{\Psi}}(\boldsymbol{a}) = \mathbb{E}_{\mathrm{X}}\left[\boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{X})\right] \ ,$$

Insightful quantities of interest

From the probability distribution of $\boldsymbol{a} \rightarrow \boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{X})$, we can derive statistical quantities of interest.

- Classical quantity of interest: The mean seismic fragility curve

$$\bar{\boldsymbol{\Psi}}(\boldsymbol{a}) = \mathbb{E}_{\mathrm{X}}\left[\boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{X})\right] \ ,$$

- Our contribution: The seismic fragility quantile curves of level $\gamma \in [\mathbf{0}, \mathbf{1}]$

$$\boldsymbol{q}_{\gamma}(\boldsymbol{a}) = \inf_{\boldsymbol{q} \in \mathbb{R}} \left\{ \mathbb{P}_{\mathrm{X}}(\boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{X}) \leq \boldsymbol{q}) \geq \gamma \right\} .$$

Insightful quantities of interest

From the probability distribution of $\boldsymbol{a} \to \boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{X})$, we can derive statistical quantities of interest.

- Classical quantity of interest: The mean seismic fragility curve

$$\bar{\boldsymbol{\Psi}}(\boldsymbol{a}) = \mathbb{E}_{\mathrm{X}}\left[\boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{X})\right] ,$$

- Our contribution: The seismic fragility quantile curves of level $\gamma \in [\mathbf{0}, \mathbf{1}]$

$$\boldsymbol{q}_{\gamma}(\boldsymbol{a}) = \inf_{\boldsymbol{q} \in \mathbb{R}} \left\{ \mathbb{P}_{\mathrm{X}}(\boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{X}) \leq \boldsymbol{q}) \geq \gamma \right\} .$$

These quantities are untractable to estimate directly using a mechanical computer code ⇒ Use of surrogate model.

Outline of the defense

1. Uncertainty Quantification
2. Surrogate modeling of computer codes using Gaussian processes
   - Seismic fragility curves estimation with Gaussian processes.
   - Uncertainty propagation
   - Application to a piping system of a French PWR
3. Global sensitivity analysis on seismic fragility curves
4. Sequential design of experiments
5. Conclusion and perspectives

Statistical model

A statistical model has to be proposed on the mechanical computer model output $\boldsymbol{z}(\boldsymbol{a}, \mathrm{x})$:

$$\boldsymbol{y}(\boldsymbol{a}, \mathrm{x}) = \boldsymbol{g}(\boldsymbol{a}, \mathrm{x}) + \boldsymbol{\varepsilon}(\boldsymbol{a}, \mathrm{x}) \ ,$$

where $\boldsymbol{\varepsilon} \sim \boldsymbol{\mathcal{N}}(0, \boldsymbol{\sigma}_{\varepsilon}^2(\boldsymbol{a}, \mathrm{x}))$ and $\boldsymbol{y}(\boldsymbol{a}, \mathrm{x}) = \log(\boldsymbol{z}(\boldsymbol{a}, \mathrm{x}))$.

Statistical model

A statistical model has to be proposed on the mechanical computer model output $\boldsymbol{z}(\boldsymbol{a}, \mathrm{x})$:

$$\boldsymbol{y}(\boldsymbol{a}, \mathrm{x}) = \boldsymbol{g}(\boldsymbol{a}, \mathrm{x}) + \boldsymbol{\varepsilon}(\boldsymbol{a}, \mathrm{x}) \ ,$$

where $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \boldsymbol{\sigma}_{\varepsilon}^2(\boldsymbol{a}, \mathrm{x}))$ and $\boldsymbol{y}(\boldsymbol{a}, \mathrm{x}) = \log(\boldsymbol{z}(\boldsymbol{a}, \mathrm{x}))$.

The seismic fragility model in this model boils down to

$$\boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{x}; \boldsymbol{g}) = \boldsymbol{\Phi}\left(\frac{\boldsymbol{g}(\boldsymbol{a}, \mathrm{x}) - \log(C)}{\boldsymbol{\sigma}_{\varepsilon}(\boldsymbol{a}, \mathrm{x})}\right) \ ,$$

where $\boldsymbol{\Phi}$ is the cdf of the standard Gaussian distribution.

Statistical model

A statistical model has to be proposed on the mechanical computer model output $\boldsymbol{z}(\boldsymbol{a}, \mathrm{x})$:

$$\boldsymbol{y}(\boldsymbol{a}, \mathrm{x}) = \boldsymbol{g}(\boldsymbol{a}, \mathrm{x}) + \boldsymbol{\varepsilon}(\boldsymbol{a}, \mathrm{x}) \ ,$$

where $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \boldsymbol{\sigma}_{\varepsilon}^2(\boldsymbol{a}, \mathrm{x}))$ and $\boldsymbol{y}(\boldsymbol{a}, \mathrm{x}) = \log(\boldsymbol{z}(\boldsymbol{a}, \mathrm{x}))$.

The seismic fragility model in this model boils down to

$$\boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{x}; \boldsymbol{g}) = \boldsymbol{\Phi}\left(\frac{\boldsymbol{g}(\boldsymbol{a}, \mathrm{x}) - \log(\boldsymbol{C})}{\boldsymbol{\sigma}_{\varepsilon}(\boldsymbol{a}, \mathrm{x})}\right) \ ,$$

where $\boldsymbol{\Phi}$ is the cdf of the standard Gaussian distribution.

$\hookrightarrow$ Objective: Build a surrogate model to learn the regression function $\boldsymbol{g}$.

Surrogate model

Main principle → Predict $\boldsymbol{y}(\boldsymbol{a}, \mathrm{x})$ using a learning dataset $(\boldsymbol{a}_i, \mathrm{x}_i, \boldsymbol{y}(\boldsymbol{a}_i, \mathrm{x}_i))_{1 \leq i \leq n}$.

[6] C. Soize and R. Ghanem. Physical systems with random uncertainties: Chaos representations with arbitrary probability measure.
SIAM Journal on Scientific Computing, 26(2):395–410, 2004

Surrogate model

Main principle → Predict $\boldsymbol{y}(\boldsymbol{a}, \mathrm{x})$ using a learning dataset $(\boldsymbol{a}_i, \mathrm{x}_i, \boldsymbol{y}(\boldsymbol{a}_i, \mathrm{x}_i))_{1 \leq i \leq n}$.

A large variety of methods available in the literature (Polynomial chaos [6], random forest, spline regression,...)

[6] C. Soize and R. Ghanem. Physical systems with random uncertainties: Chaos representations with arbitrary probability measure.
SIAM Journal on Scientific Computing, 26(2):395–410, 2004

Surrogate model

Main principle → Predict $\boldsymbol{y}(\boldsymbol{a}, \mathrm{x})$ using a learning dataset $(\boldsymbol{a}_i, \mathrm{x}_i, \boldsymbol{y}(\boldsymbol{a}_i, \mathrm{x}_i))_{1 \leq i \leq n}$.

A large variety of methods available in the literature (Polynomial chaos [6], random forest, spline regression,...)

Gaussian process (GP) regression has the main advantage to propose an uncertainty on its prediction.

[6] C. Soize and R. Ghanem. Physical systems with random uncertainties: Chaos representations with arbitrary probability measure.
SIAM Journal on Scientific Computing, 26(2):395–410, 2004

Gaussian process regression

The uncertainty on the regression function $\boldsymbol{g}$ is modeled by a Gaussian process $\boldsymbol{G}$ with a mean function $\boldsymbol{m}$ and a covariance function $\boldsymbol{\Sigma_{\theta}}$. The statistical model becomes

$$\boldsymbol{Y}(\boldsymbol{a}, \mathrm{x}) = \boldsymbol{G}(\boldsymbol{a}, \mathrm{x}) + \boldsymbol{\varepsilon}(\boldsymbol{a}, \mathrm{x}) \; .$$

Gaussian process regression

The uncertainty on the regression function $\boldsymbol{g}$ is modeled by a Gaussian process $\boldsymbol{G}$ with a mean function $\boldsymbol{m}$ and a covariance function $\boldsymbol{\Sigma_\theta}$. The statistical model becomes

$$\boldsymbol{Y}(\boldsymbol{a}, \mathrm{x}) = \boldsymbol{G}(\boldsymbol{a}, \mathrm{x}) + \boldsymbol{\varepsilon}(\boldsymbol{a}, \mathrm{x}) \; .$$

Given a learning set $\boldsymbol{\mathcal{D}_n} = (\boldsymbol{a}_i, \mathrm{x}_i, \boldsymbol{y}(\boldsymbol{a}_i, \mathrm{x}_i))_{1 \leq i \leq n}$, we can obtain the posterior distribution

$$(\boldsymbol{G}(\boldsymbol{a}, \mathrm{x}) | \boldsymbol{\mathcal{D}_n}) \sim \boldsymbol{\mathcal{N}}(\widehat{\boldsymbol{G}}_n(\boldsymbol{a}, \mathrm{x}), \widehat{\boldsymbol{\sigma}}_n(\boldsymbol{a}, \mathrm{x})^2) \; ,$$

Gaussian process regression

The uncertainty on the regression function $\boldsymbol{g}$ is modeled by a Gaussian process $\boldsymbol{G}$ with a mean function $\boldsymbol{m}$ and a covariance function $\boldsymbol{\Sigma_\theta}$. The statistical model becomes

$$\boldsymbol{Y}(\boldsymbol{a}, \mathrm{x}) = \boldsymbol{G}(\boldsymbol{a}, \mathrm{x}) + \boldsymbol{\varepsilon}(\boldsymbol{a}, \mathrm{x}) \;.$$

Given a learning set $\boldsymbol{\mathcal{D}_n} = (\boldsymbol{a}_i, \mathrm{x}_i, \boldsymbol{y}(\boldsymbol{a}_i, \mathrm{x}_i))_{1 \leq i \leq n}$, we can obtain the posterior distribution

$$(\boldsymbol{G}(\boldsymbol{a}, \mathrm{x}) | \boldsymbol{\mathcal{D}_n}) \sim \boldsymbol{\mathcal{N}}(\widehat{\boldsymbol{G}}_n(\boldsymbol{a}, \mathrm{x}), \widehat{\boldsymbol{\sigma}}_n(\boldsymbol{a}, \mathrm{x})^2) \;,$$

$\hookrightarrow$ A posterior distribution is obtained on the regression function $\boldsymbol{g}$.

Seismic fragility curve estimator

Given a learning dataset $\boldsymbol{\mathcal{D}}_n$, we can estimate the seismic fragility curve using the posterior mean.

$$\boldsymbol{\Psi}^{(1)}(\boldsymbol{a}, \mathrm{x}) = \mathbb{E}_{G}\left[\boldsymbol{\Psi}(\boldsymbol{a}, \mathrm{x}; G) | \boldsymbol{\mathcal{D}}_n\right] \ ,$$

Seismic fragility curve estimator

Given a learning dataset $\mathcal{D}_n$, we can estimate the seismic fragility curve using the posterior mean.

$$\Psi^{(1)}(a, \mathrm{x}) = \mathbb{E}_G \left[\Psi(a, \mathrm{x}; G) | \mathcal{D}_n\right] \ ,$$

$$\Psi^{(1)}(a, \mathrm{x}) = \Phi\left(\frac{\widehat{G}_n(a, \mathrm{x}) - \log(C)}{\sigma_n(a, \mathrm{x})}\right) \ ,$$

where $\sigma_n(a, \mathrm{x})^2 = \widehat{\sigma}_n(a, \mathrm{x})^2 + \sigma_\varepsilon(a, \mathrm{x})^2$.

$\Psi^{(1)}$ is deterministic.

Surrogate model uncertainty propagation

The advantage of the GP surrogate model is the possibility to propagate its uncertainty on the seismic fragility curve

$$\Psi^{(2)}(a, \mathrm{x}) = \Phi\left(\frac{G_n(a, \mathrm{x}) - \log(C)}{\sigma_\varepsilon(a, \mathrm{x})}\right) ,$$

where $G_n \overset{\mathcal{L}}{=} (G|\mathcal{D}_n)$.

Surrogate model uncertainty propagation

The advantage of the GP surrogate model is the possibility to propagate its uncertainty on the seismic fragility curve

$$\boldsymbol{\Psi}^{(2)}(\boldsymbol{a}, \mathrm{x}) = \boldsymbol{\Phi}\left(\frac{\boldsymbol{G}_n(\boldsymbol{a}, \mathrm{x}) - \log(\boldsymbol{C})}{\boldsymbol{\sigma}_{\varepsilon}(\boldsymbol{a}, \mathrm{x})}\right),$$

where $\boldsymbol{G}_n \stackrel{\mathcal{L}}{=} (\boldsymbol{G}|\boldsymbol{\mathcal{D}}_n)$.

$\boldsymbol{\Psi}^{(2)}$ is random (due to $\boldsymbol{G}_n$).

- **Plug-in estimator.** Consider a Monte-Carlo sample $(\mathbf{X}_j)_{1\leq j\leq m}$.

$$q_{\gamma}^{(1)}(a) = \inf_{q\in\mathbb{R}} \left\{ \frac{1}{m} \sum_{j=1}^{m} \mathbb{1}_{(\Psi^{(1)}(a,\mathbf{X}_j)\leq q)} \geq \gamma \right\}.$$

Seismic fragility quantile curve estimation

- Plug-in estimator. Consider a Monte-Carlo sample $(\mathbf{X}_j)_{1\leq j\leq m}$.

$$q_\gamma^{(1)}(a) = \inf_{q\in\mathbb{R}} \left\{ \frac{1}{m}\sum_{j=1}^{m} \mathbb{1}_{(\Psi^{(1)}(a,\mathbf{X}_j)\leq q)} \geq \gamma \right\}.$$

- Bi-level seismic fragility quantile. Consider a sample $(\Psi_p^{(2)})_{1\leq p\leq P}$ of $\Psi^{(2)}$.

$$q_{\gamma_G}^{(2)}(a,\mathbf{X}) = \inf_{q\in\mathbb{R}} \left\{ \frac{1}{P}\sum_{p=1}^{P} \mathbb{1}_{(\Psi_p^{(2)}(a,\mathbf{X})\leq q)} \geq \gamma_G \right\}.$$

$\hookrightarrow$ $q_{\gamma_G}^{(2)}$ is a quantile w.r.t. the GP surrogate posterior distribution.

Seismic fragility quantile curve estimation

- Plug-in estimator. Consider a Monte-Carlo sample $(\mathrm{X}_j)_{1\leq j\leq m}$.

$$q_{\gamma}^{(1)}(a) = \inf_{q\in\mathbb{R}} \left\{ \frac{1}{m} \sum_{j=1}^{m} \mathbb{1}_{(\Psi^{(1)}(a,\mathrm{X}_j)\leq q)} \geq \gamma \right\}.$$

- Bi-level seismic fragility quantile. Consider a sample $(\Psi_p^{(2)})_{1\leq p\leq P}$ of $\Psi^{(2)}$.

$$q_{\gamma_G}^{(2)}(a,\mathrm{X}) = \inf_{q\in\mathbb{R}} \left\{ \frac{1}{P} \sum_{p=1}^{P} \mathbb{1}_{(\Psi_p^{(2)}(a,\mathrm{X})\leq q)} \geq \gamma_G \right\}.$$

$\hookrightarrow$ $q_{\gamma_G}^{(2)}$ is a quantile w.r.t. the GP surrogate posterior distribution.

$$q_{\gamma_G,\gamma}^{(2)}(a) = \inf_{q\in\mathbb{R}} \left\{ \frac{1}{m} \sum_{j=1}^{m} \mathbb{1}_{(q_{\gamma_G}^{(2)}(a,\mathrm{X}_j)\leq q)} \geq \gamma \right\}$$

$\hookrightarrow$ $q_{\gamma_G,\gamma}^{(2)}$ is a bi-level seismic fragility quantile curve (GP uncertainty + epistemic uncertainties).

Industrial use case: The *Alimentation de Secours Générale* (ASG) piping system, equipping French nuclear power plants.

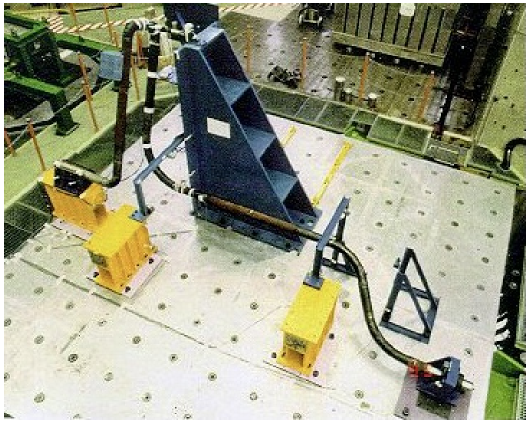


Figure: A view of the ASG piping model on the Azalée shaking table at CEA Saclay

[7] F. Touboul, P. Sollogoub, and N. Blay. Seismic behaviour of piping systems with and without defects: experimental and numerical evaluations.
Nuclear Engineering and Design, 192(2):243–260, 1999

Industrial use case: ASG piping system

Industrial use case: The *Alimentation de Secours Générale* (ASG) piping system, equipping French nuclear power plants.

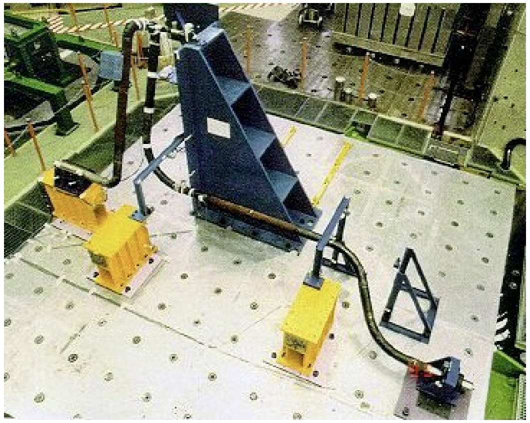


Figure: A view of the ASG piping model on the Azalée shaking table at CEA Saclay

- CAST3M finite element code simulating the dynamic behavior of the piping system validated on the basis of experimental tests [7];

[7] F. Touboul, P. Sollogoub, and N. Blay. Seismic behaviour of piping systems with and without defects: experimental and numerical evaluations.
Nuclear Engineering and Design, 192(2):243–260, 1999

Industrial use case: ASG piping system

Industrial use case: The *Alimentation de Secours Générale* (ASG) piping system, equipping French nuclear power plants.

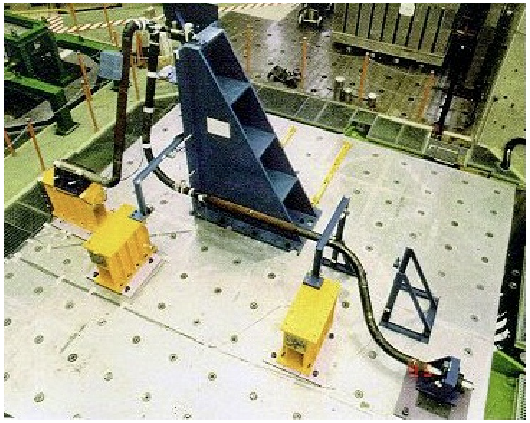


Figure: A view of the ASG piping model on the Azalée shaking table at CEA Saclay

- CAST3M finite element code simulating the dynamic behavior of the piping system validated on the basis of experimental tests [7];
- Ten structural parameters (Boundary conditions, mechanical characteristics);

[7] F. Touboul, P. Sollogoub, and N. Blay. Seismic behaviour of piping systems with and without defects: experimental and numerical evaluations. *Nuclear Engineering and Design*, 192(2):243–260, 1999

Industrial use case: ASG piping system

Industrial use case: The *Alimentation de Secours Générale* (ASG) piping system, equipping French nuclear power plants.

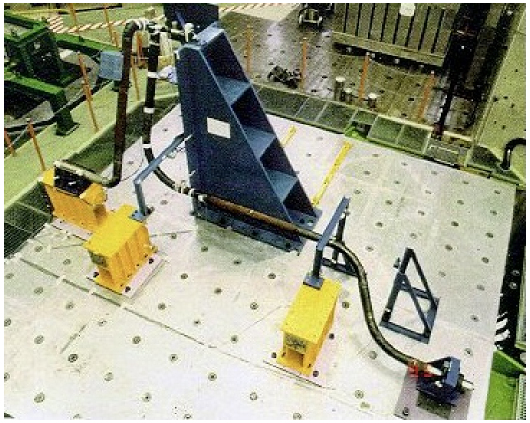


Figure: A view of the ASG piping model on the Azalée shaking table at CEA Saclay

- CAST3M finite element code simulating the dynamic behavior of the piping system validated on the basis of experimental tests [7];
- Ten structural parameters (Boundary conditions, mechanical characteristics);
- A quantity of interest for the reliability of the structure: the out-of-plane rotation of a pipe elbow;

[7] F. Touboul, P. Sollogoub, and N. Blay. Seismic behaviour of piping systems with and without defects: experimental and numerical evaluations. *Nuclear Engineering and Design*, 192(2):243–260, 1999

Industrial use case: ASG piping system

Industrial use case: The *Alimentation de Secours Générale* (ASG) piping system, equipping French nuclear power plants.

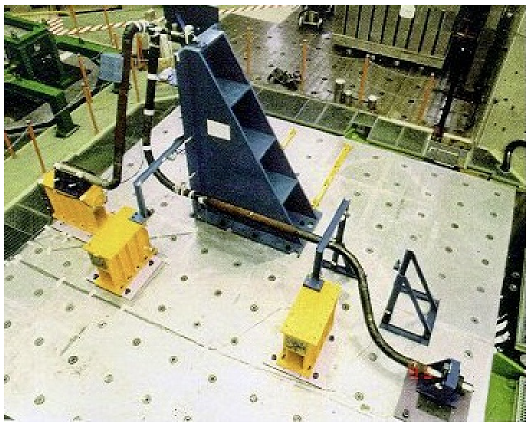


Figure: A view of the ASG piping model on the Azalée shaking table at CEA Saclay

- CAST3M finite element code simulating the dynamic behavior of the piping system validated on the basis of experimental tests [7];
- Ten structural parameters (Boundary conditions, mechanical characteristics);
- A quantity of interest for the reliability of the structure: the out-of-plane rotation of a pipe elbow;
- Computational cost of 1 CAST3M call ≈ 1 minute ⇒ 100 days of computation time for uncertainty propagation.

[7] F. Touboul, P. Sollogoub, and N. Blay. Seismic behaviour of piping systems with and without defects: experimental and numerical evaluations. *Nuclear Engineering and Design*, 192(2):243–260, 1999

Sources of epistemic uncertainties

The epistemic uncertainties on the ASG piping system have two main sources:

Sources of epistemic uncertainties

The epistemic uncertainties on the ASG piping system have two main sources:

- ↪ Material parameters: Young modulus, elasticity limit, hardening module and modal damping ratio;

Sources of epistemic uncertainties

The epistemic uncertainties on the ASG piping system have two main sources:

- ↪ Material parameters: Young modulus, elasticity limit, hardening module and modal damping ratio;
- ↪ Boundary conditions: Translation and rotation stiffness at the clamped end and in the guide.

Structural parameters

Variable number	*Variable*
1	E: Young modulus
2	Sy: Elasticity limit
3	H: Hardening module
4	b: Modal damping ratio
5	RPY151: Rotation stiffness for the P151 guide in Y direction
6	RPX29: Rotation stiffness for the P29 clamped end in X direction
7	RPY29: Rotation stiffness for the P29 clamped end in Y direction
8	TPX29: Translation stiffness for the P29 clamped end in X direction
9	TPY29: Translation stiffness for the P29 clamped end in Y direction
10	TPZ29: Translation stiffness for the P29 clamped end in Z direction

↪ Material parameters

↪ Boundary conditions parameters

ASG use case settings

- The seismic intensity measure chosen is the spectral acceleration at pulsation $f = 5$ Hz and damping ratio $\xi = 0.01$

$$a = \max_{t \in [0,T]} (2\pi f)^2 |\ell(t)| \,,$$

where ℓ is the displacement of a single d.o.f. oscillator;

[8] A. Marrel, B. Iooss, and V. Chabridon. The icscream methodology: Identification of penalizing configurations in computer experiments using screening and metamodel—applications in thermal-hydraulics. *Nuclear Science and Engineering*, 196(3):301–321, 2022

[9] S. Rezaeian and A. Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 39(10):1155–1180, 2010

ASG use case settings

- The seismic intensity measure chosen is the spectral acceleration at pulsation $f = 5$ Hz and damping ratio $\xi = 0.01$

$$a = \max_{t \in [0,T]} (2\pi f)^2 |\ell(t)| \ ,$$

where ℓ is the displacement of a single d.o.f. oscillator;

- A screening step using HSIC-based tests [8] is performed to select the 6 most influential mechanical parameters, gathered in the vector $\mathbf{X}$;

[8] A. Marrel, B. Iooss, and V. Chabridon. The icscream methodology: Identification of penalizing configurations in computer experiments using screening and metamodel—applications in thermal-hydraulics. *Nuclear Science and Engineering*, 196(3):301–321, 2022

[9] S. Rezaeian and A. Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 39(10):1155–1180, 2010

ASG use case settings

- The seismic intensity measure chosen is the spectral acceleration at pulsation $f = 5$ Hz and damping ratio $\xi = 0.01$

$$a = \max_{t \in [0,T]} (2\pi f)^2 |\ell(t)| \ ,$$

where ℓ is the displacement of a single d.o.f. oscillator;

- A screening step using HSIC-based tests [8] is performed to select the 6 most influential mechanical parameters, gathered in the vector $\mathbf{X}$;

- Synthetic ground motions are generated using a stochastic simulator [9] and filtered by a one d.o.f oscillator modeling a fictitious building;

---

[8] A. Marrel, B. Iooss, and V. Chabridon. The icscream methodology: Identification of penalizing configurations in computer experiments using screening and metamodel—applications in thermal-hydraulics. *Nuclear Science and Engineering*, 196(3):301–321, 2022

[9] S. Rezaeian and A. Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 39(10):1155–1180, 2010

ASG use case settings

- The seismic intensity measure chosen is the spectral acceleration at pulsation $f = 5$ Hz and damping ratio $\xi = 0.01$

$$a = \max_{t \in [0,T]} (2\pi f)^2 |\ell(t)| \,,$$

where ℓ is the displacement of a single d.o.f. oscillator;

- A screening step using HSIC-based tests [8] is performed to select the 6 most influential mechanical parameters, gathered in the vector $\mathbf{X}$;

- Synthetic ground motions are generated using a stochastic simulator [9] and filtered by a one d.o.f oscillator modeling a fictitious building;

- 500 nonlinear mechanical simulations are carried out using CAST3M finite element code.

[8] A. Marrel, B. Iooss, and V. Chabridon. The icscream methodology: Identification of penalizing configurations in computer experiments using screening and metamodel—applications in thermal-hydraulics. *Nuclear Science and Engineering*, 196(3):301–321, 2022

[9] S. Rezaeian and A. Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 39(10):1155–1180, 2010

Gaussian process surrogates

Two Gaussian processes are proposed as surrogates for the mechanical computer code.

They are both zero mean with a tensorized anisotropic Matérn 5/2 covariance function.

[10] A. P. Kyprioti and A. A. Taflanidis. Kriging metamodeling for seismic response distribution estimation. *Earthquake Engineering & Structural Dynamics*, 50(13):3550–3576, 2021

Gaussian process surrogates

Two Gaussian processes are proposed as surrogates for the mechanical computer code.

They are both zero mean with a tensorized anisotropic Matérn 5/2 covariance function.

- $\hookrightarrow$ Homoskedastic: $\boldsymbol{\sigma}_{\varepsilon}(\boldsymbol{a}, \mathbf{x}) = \boldsymbol{\sigma}_{\varepsilon}$;

[10] A. P. Kyprioti and A. A. Taflanidis. Kriging metamodeling for seismic response distribution estimation. *Earthquake Engineering & Structural Dynamics*, 50(13):3550–3576, 2021

Gaussian process surrogates

Two Gaussian processes are proposed as surrogates for the mechanical computer code.

They are both zero mean with a tensorized anisotropic Matérn 5/2 covariance function.

- $\hookrightarrow$ Homoskedastic: $\boldsymbol{\sigma}_\varepsilon(\boldsymbol{a}, \mathrm{x}) = \boldsymbol{\sigma}_\varepsilon$;
- $\hookrightarrow$ Heteroskedastic: $\boldsymbol{\sigma}_\varepsilon(\boldsymbol{a}, \mathrm{x}) = \max(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{a}, \boldsymbol{\beta}_2)$ (Parametric model adapted from Kyprioti and Taflanidis [2021][10]).

[10] A. P. Kyprioti and A. A. Taflanidis. Kriging metamodeling for seismic response distribution estimation. *Earthquake Engineering & Structural Dynamics*, 50(13):3550–3576, 2021

Gaussian process surrogates

Two Gaussian processes are proposed as surrogates for the mechanical computer code.

They are both zero mean with a tensorized anisotropic Matérn 5/2 covariance function.

- $\hookrightarrow$ Homoskedastic: $\boldsymbol{\sigma}_\varepsilon(\boldsymbol{a}, \mathrm{x}) = \boldsymbol{\sigma}_\varepsilon$;
- $\hookrightarrow$ Heteroskedastic: $\boldsymbol{\sigma}_\varepsilon(\boldsymbol{a}, \mathrm{x}) = \mathbf{max}(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{a}, \boldsymbol{\beta}_2)$ (Parametric model adapted from Kyprioti and Taflanidis [2021][10]).

Heteroskedastic model motivation $\rightarrow$ Mechanical nonlinearities for high intensity seismic ground motions.

[10] A. P. Kyprioti and A. A. Taflanidis. Kriging metamodeling for seismic response distribution estimation. *Earthquake Engineering & Structural Dynamics*, 50(13):3550–3576, 2021

Performance evaluation

Learning sample size	100	200	300	400	500
Homoskedastic	**0.844**	**0.860**	**0.853**	**0.870**	**0.867**
Heteroskedastic	**0.842**	**0.860**	**0.849**	**0.872**	**0.875**

Table: Q^2 numerical values estimated by leave-one-out.

Performance evaluation

Learning sample size	100	200	300	400	500
Homoskedastic	**0.844**	**0.860**	**0.853**	**0.870**	**0.867**
Heteroskedastic	**0.842**	**0.860**	**0.849**	**0.872**	**0.875**

Table: $\boldsymbol{Q^2}$ numerical values estimated by leave-one-out.

$\hookrightarrow$ Same predictive performance for the two models.

Coverage probability

Credible intervals of level $\boldsymbol{\alpha} \in [\mathbf{0}, \mathbf{1}]$

$$\mathbf{CI}_{\alpha}(\boldsymbol{a}, \mathrm{x}) = \left[\widehat{\boldsymbol{G}}_n(\boldsymbol{a}, \mathrm{x}) - \boldsymbol{q}_{1-\frac{\alpha}{2}} \boldsymbol{\sigma}_n(\boldsymbol{a}, \mathrm{x}), \widehat{\boldsymbol{G}}_n(\boldsymbol{a}, \mathrm{x}) + \boldsymbol{q}_{1-\frac{\alpha}{2}} \boldsymbol{\sigma}_n(\boldsymbol{a}, \mathrm{x})\right] ,$$

$\boldsymbol{q}_{\gamma}$ is the γ-level quantile of the standard Gaussian distribution. $\boldsymbol{\sigma}_n(\boldsymbol{a}, \mathrm{x})^2 = \widehat{\boldsymbol{\sigma}}_n(\boldsymbol{a}, \mathrm{x})^2 + \boldsymbol{\sigma}_{\varepsilon}^{\mathbf{2}}$.

Coverage probability

Credible intervals of level $\boldsymbol{\alpha} \in [\mathbf{0}, \mathbf{1}]$

$$\mathbf{CI}_\alpha(\boldsymbol{a}, \mathrm{x}) = \left[\widehat{\boldsymbol{G}}_n(\boldsymbol{a}, \mathrm{x}) - \boldsymbol{q}_{1-\frac{\alpha}{2}}\boldsymbol{\sigma}_n(\boldsymbol{a}, \mathrm{x}), \widehat{\boldsymbol{G}}_n(\boldsymbol{a}, \mathrm{x}) + \boldsymbol{q}_{1-\frac{\alpha}{2}}\boldsymbol{\sigma}_n(\boldsymbol{a}, \mathrm{x})\right] \ ,$$

$\boldsymbol{q}_\gamma$ is the γ-level quantile of the standard Gaussian distribution. $\boldsymbol{\sigma}_n(\boldsymbol{a}, \mathrm{x})^2 = \widehat{\boldsymbol{\sigma}}_n(\boldsymbol{a}, \mathrm{x})^2 + \boldsymbol{\sigma}_\varepsilon^2$.

Coverage probability of level $\boldsymbol{\alpha}$

$$\mathbf{CP}_\alpha = \mathbb{P}(\boldsymbol{y}(\boldsymbol{A}, \mathrm{X}) \in \mathbf{CI}_\alpha(\boldsymbol{A}, \mathrm{X}))\ .$$

$\hookrightarrow$ Similarity measure between the data distribution and the GP posterior distribution.

Coverage probability

Credible intervals of level $\alpha \in [0, 1]$

$$\mathrm{CI}_\alpha(a, \mathrm{x}) = \left[\widehat{G}_n(a, \mathrm{x}) - q_{1-\frac{\alpha}{2}}\sigma_n(a, \mathrm{x}), \widehat{G}_n(a, \mathrm{x}) + q_{1-\frac{\alpha}{2}}\sigma_n(a, \mathrm{x})\right] ,$$

q_γ is the γ-level quantile of the standard Gaussian distribution. $\sigma_n(a, \mathrm{x})^2 = \widehat{\sigma}_n(a, \mathrm{x})^2 + \sigma_\varepsilon^2$.

Coverage probability of level α

$$\mathrm{CP}_\alpha = \mathbb{P}(y(A, \mathrm{X}) \in \mathrm{CI}_\alpha(A, \mathrm{X})) .$$

$\hookrightarrow$ Similarity measure between the data distribution and the GP posterior distribution.

Empirical estimator on a test dataset $(a'_i, \mathrm{x}'_i, y(a'_i, \mathrm{x}'_i))_{1 \leq i \leq n'}$ or by cross-validation.

$$\widehat{\mathrm{CP}_\alpha} = \frac{1}{n'} \sum_{i=1}^{n'} \mathbb{1}_{y(a'_i, \mathrm{x}'_i) \in \mathrm{CI}_\alpha(a'_i, \mathrm{x}'_i)} .$$

Performance evaluation

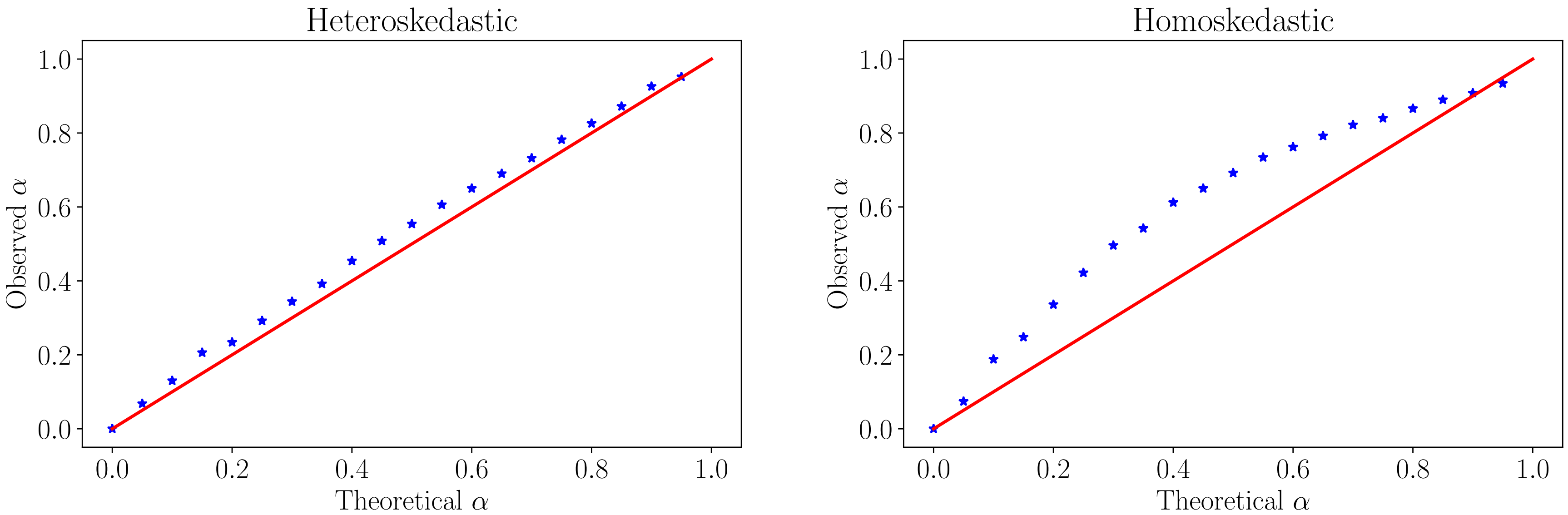


Figure: Leave-one-out estimation of the coverage probabilities.

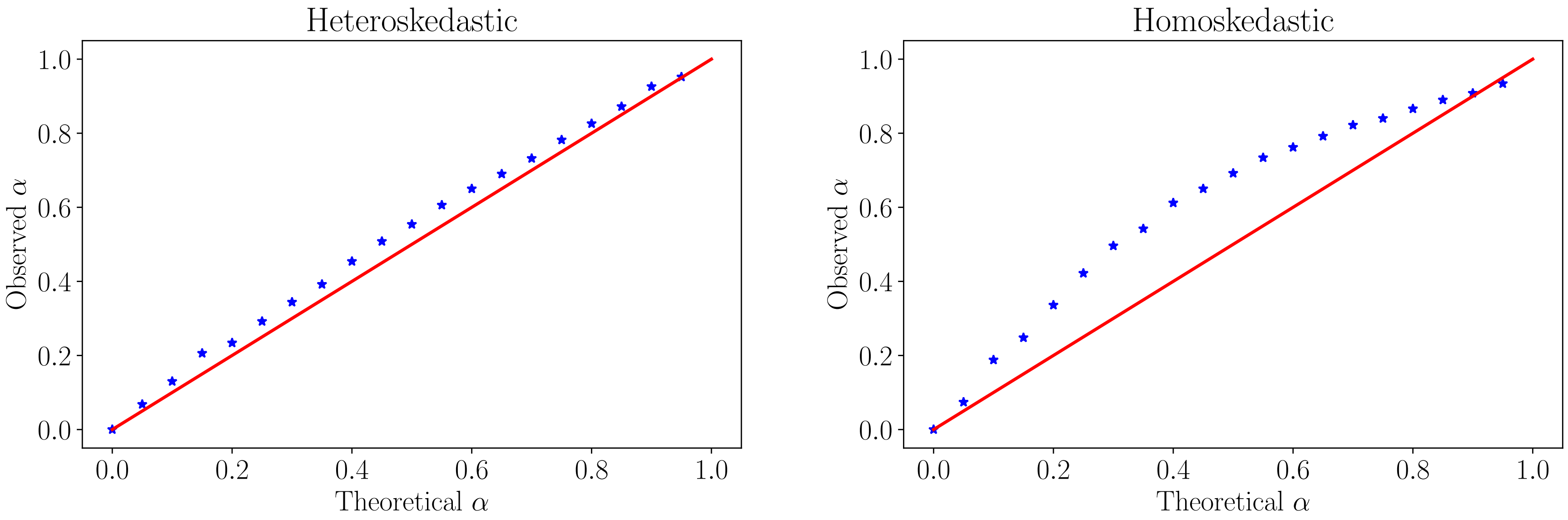
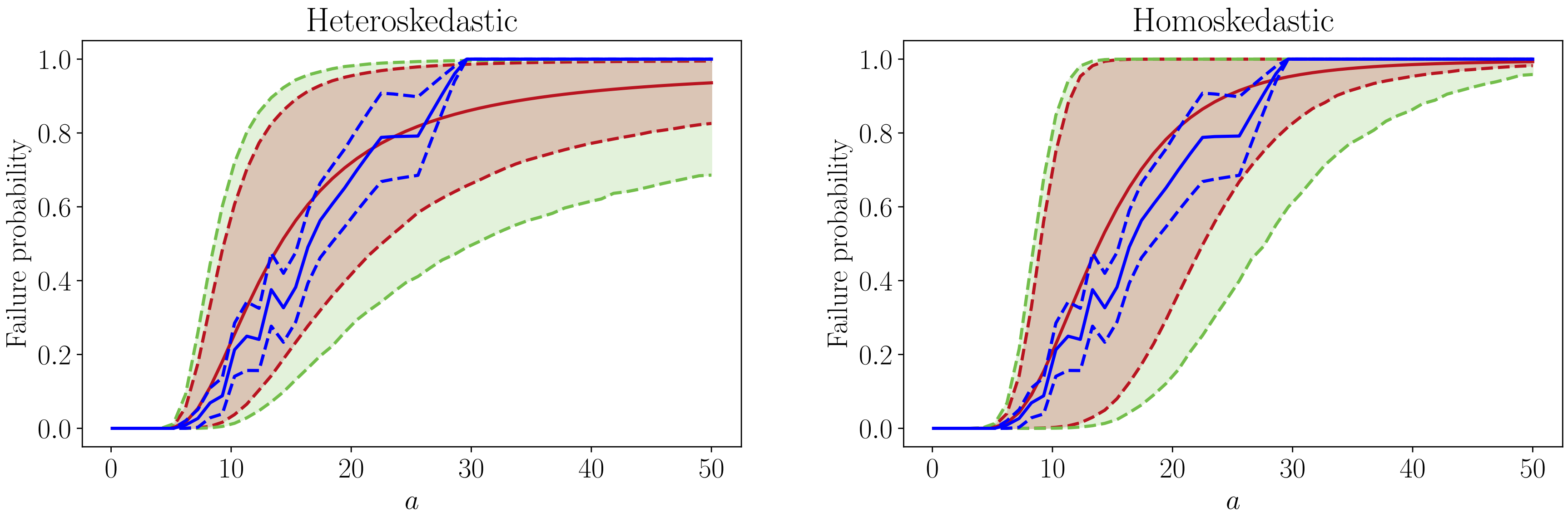


Figure: Leave-one-out estimation of the coverage probabilities.

↪ Better coverage probabilities with the heteroskedastic model.

Uncertainty propagation



- $n = 500$ CAST3M computations
- Plug-in quantile curve estimator ($\gamma = 0.1$ and $\gamma = 0.9$)
- Bi-level quantile curves estimator ($\gamma = \gamma_G = 0.1$ and $\gamma = \gamma_G = 0.9$)
- Nonparametric estimation of the mean curve (2000 CAST3M computations)

Outline of the defense

Motivation

Sensitivity analysis (SA) aims at studying how the uncertainty on the model output can be apportioned to the different uncertainties sources of the model inputs.

Motivation

Sensitivity analysis (SA) aims at studying how the uncertainty on the model output can be apportioned to the different uncertainties sources of the model inputs.

Global sensitivity analysis (GSA) takes into account the overall uncertainty of the model inputs.

Motivation

Sensitivity analysis (SA) aims at studying how the uncertainty on the model output can be apportioned to the different uncertainties sources of the model inputs.

Global sensitivity analysis (GSA) takes into account the overall uncertainty of the model inputs.

Model inputs $\rightarrow$ Mechanical parameters of the structure

Model output $\rightarrow$ Seismic fragility curve

Aggregated Sobol' indices

The aggregated Sobol' indices are a natural extension of Sobol' indices to a functional output [11].

[11] B. Iooss and L. Le Gratiet. Uncertainty and sensitivity analysis of functional risk curves based on Gaussian processes. *Reliability Engineering & System Safety*, 187:58–66, 2019

Aggregated Sobol' indices

The aggregated Sobol' indices are a natural extension of Sobol' indices to a functional output [11].

Using the notation $\mathbf{X} = (\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(p)})$

$$D = \int_{\mathcal{A}} \mathrm{Var}(\Psi(a, \mathrm{X}))dh(a)$$

$$S_i^{\mathrm{FC}} = \frac{1}{D} \int_{\mathcal{A}} \mathrm{Var}\left(\mathbb{E}[\Psi(a, \mathrm{X})|\mathbf{X}^{(i)}]\right) dh(a)$$

is the first-order aggregated Sobol' index of $\mathbf{X}^{(i)}$.

[11] B. Iooss and L. Le Gratiet. Uncertainty and sensitivity analysis of functional risk curves based on Gaussian processes. *Reliability Engineering & System Safety*, 187:58–66, 2019

The aggregated Sobol' indices are a natural extension of Sobol' indices to a functional output [11].

Using the notation $\mathbf{X} = (\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(p)})$

$$D = \int_{\mathcal{A}} \mathrm{Var}(\Psi(a, \mathrm{X}))dh(a)$$

$$S_i^{\mathrm{FC}} = \frac{1}{D} \int_{\mathcal{A}} \mathrm{Var}\left(\mathbb{E}[\Psi(a, \mathrm{X})|\mathbf{X}^{(i)}]\right) dh(a)$$

is the first-order aggregated Sobol' index of $\mathbf{X}^{(i)}$.

The probability measure $h \rightarrow$ Choice of the practicioner.

[11] B. Iooss and L. Le Gratiet. Uncertainty and sensitivity analysis of functional risk curves based on Gaussian processes. *Reliability Engineering & System Safety*, 187:58–66, 2019

ANOVA decomposition

Our contribution: ANOVA decomposition of the aggregated Sobol′ indices.

Proposition

The aggregated Sobol′ indices follow a ANOVA decomposition

$$\sum_{u \subseteq \{1,\ldots,p\}} S_u^{FC} = 1 \, ,$$

where $\mathbf{X}^{(u)} = (\mathbf{X}^{(j)})_{j \in u}$ *and*
$S_u^{FC} = \sum_{v \subseteq u} (-1)^{|u|-|v|} \int_{\mathcal{A}} \mathrm{Var}\left(\mathbb{E}[\Psi(a, \mathrm{X}) | \mathrm{X}^{(v)}]\right) dh(a).$

ANOVA decomposition

Our contribution: ANOVA decomposition of the aggregated Sobol′ indices.

Proposition

The aggregated Sobol′ indices follow a ANOVA decomposition

$$\sum_{u\subseteq\{1,\dots,p\}} S_u^{FC} = 1 ,$$

where $\mathbf{X}^{(u)} = (\mathbf{X}^{(j)})_{j\in u}$ *and*
$S_u^{FC} = \sum_{v\subseteq u}(-1)^{|u|-|v|}\int_{\mathcal{A}} \mathrm{Var}\left(\mathbb{E}[\Psi(a,\mathrm{X})|\mathrm{X}^{(v)}]\right) dh(a)$.

↪ The total-order aggregated Sobol′ indices are thus well-defined

$$T_i^{FC} = 1 - \frac{1}{D}\int_{\mathcal{A}} \mathrm{Var}\left(\mathbb{E}[\Psi(a,\mathrm{X})|\mathrm{X}^{(-i)}]\right) dh(a) ,$$

where $\mathbf{X}^{(-i)} = (\mathbf{X}^{(j)})_{j\neq i}$.

Kernel methods

Recently, global sensitivity indices based on kernels have been proposed [12].

[12] J. Barr and H. Rabitz. A generalized kernel method for global sensitivity analysis. *SIAM/ASA Journal on Uncertainty Quantification*, 10(1):27–54, 2022

[13] S. Da Veiga, F. Gamboa, B. Iooss, and C. Prieur. *Basics and Trends in Sensitivity Analysis.* Society for Industrial and Applied Mathematics, 2021

Recently, global sensitivity indices based on kernels have been proposed [12].

Theoretical guarantees $\rightarrow$ ANOVA decomposition [13].

[12] J. Barr and H. Rabitz. A generalized kernel method for global sensitivity analysis. *SIAM/ASA Journal on Uncertainty Quantification*, 10(1):27–54, 2022

[13] S. Da Veiga, F. Gamboa, B. Iooss, and C. Prieur. *Basics and Trends in Sensitivity Analysis.* Society for Industrial and Applied Mathematics, 2021

Kernel methods

Recently, global sensitivity indices based on kernels have been proposed [12].

Theoretical guarantees → ANOVA decomposition [13].

Motivation: Kernels methods are adapted to complex data types.

↪ Definition of kernel based global sensitivity indices tailored for seismic fragility curves.

[12] J. Barr and H. Rabitz. A generalized kernel method for global sensitivity analysis. *SIAM/ASA Journal on Uncertainty Quantification*, 10(1):27–54, 2022

[13] S. Da Veiga, F. Gamboa, B. Iooss, and C. Prieur. *Basics and Trends in Sensitivity Analysis.* Society for Industrial and Applied Mathematics, 2021

MMD distance

Consider a Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}$ with reproducing kernel $k : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ and dot product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$.

The kernel mean embedding $m_\mu \in \mathcal{H}$ of a probability measure μ is defined by

$$m_\mu = \int_{\mathcal{X}} k(\mathrm{x}, \cdot) d\mu(\mathrm{x}) \; .$$

[14] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(25):723–773, 2012

MMD distance

Consider a Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}$ with reproducing kernel $k : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ and dot product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$.

The kernel mean embedding $m_\mu \in \mathcal{H}$ of a probability measure μ is defined by

$$m_\mu = \int_{\mathcal{X}} k(\mathrm{x}, \cdot) d\mu(\mathrm{x}) \, .$$

The Maximum Mean Discrepancy (MMD) distance between probability measures is defined by

$$\mathrm{MMD}(\mu, \nu) = \| m_\mu - m_\nu \|_{\mathcal{H}} \, .$$

[14] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(25):723–773, 2012

MMD distance

Consider a Reproducing Kernel Hilbert Space (RKHS) $\mathcal{H}$ with reproducing kernel $k : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ and dot product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$.

The kernel mean embedding $m_\mu \in \mathcal{H}$ of a probability measure μ is defined by

$$m_\mu = \int_{\mathcal{X}} k(\mathrm{x}, \cdot) d\mu(\mathrm{x}) \, .$$

The Maximum Mean Discrepancy (MMD) distance between probability measures is defined by

$$\mathrm{MMD}(\mu, \nu) = \|m_\mu - m_\nu\|_{\mathcal{H}} \, .$$

Practical expression of the MMD [14]

$$\mathrm{MMD}^2(\mu, \nu) = \mathbb{E}_{U,U' \sim \mu}[k(U, U')] + \mathbb{E}_{V,V' \sim \nu}[k(V, V')] - 2\mathbb{E}_{U \sim \mu, V \sim \nu}[k(U, V)] \, .$$

[14] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(25):723–773, 2012

β^k indices

The β^k indices are generalized global sensitivity indices [15] defined using Maximum Mean Discrepancy (MMD) distance.

[15] S. Da Veiga. Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, 85(7):1283–1305, 2015

[16] S. Da Veiga, F. Gamboa, B. Iooss, and C. Prieur. *Basics and Trends in Sensitivity Analysis.* Society for Industrial and Applied Mathematics, 2021

β^k indices

The β^k indices are generalized global sensitivity indices [15] defined using Maximum Mean Discrepancy (MMD) distance.

The first-order β^k_i and total-order β^k_{-i} indices of variable $\mathbf{X}^{(i)}$ are defined by:

$$\beta^k_i = \frac{\mathbb{E}_{\mathbf{X}^{(i)}}\left[\mathrm{MMD}^2\left(\mathbb{P}_{\Psi}, \mathbb{P}_{\Psi|\mathbf{X}^{(i)}}\right)\right]}{\mathrm{MMD}^2_{\mathrm{tot}}},$$

$$\beta^k_{-i} = 1 - \frac{\mathbb{E}_{\mathbf{X}^{(-i)}}\left[\mathrm{MMD}^2\left(\mathbb{P}_{\Psi}, \mathbb{P}_{\Psi|\mathbf{X}^{(-i)}}\right)\right]}{\mathrm{MMD}^2_{\mathrm{tot}}},$$

$$\mathrm{MMD}^2_{\mathrm{tot}} = \mathbb{E}[k(\Psi(\cdot,\mathbf{X}), \Psi(\cdot,\mathbf{X}))] - \mathbb{E}_{\mathbf{X},\mathbf{X}'}[k(\Psi(\cdot,\mathbf{X}), \Psi(\cdot,\mathbf{X}'))].$$

$\hookrightarrow$ ANOVA decomposition for the β^k-indices [16].

---

[15] S. Da Veiga. Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, 85(7):1283–1305, 2015

[16] S. Da Veiga, F. Gamboa, B. Iooss, and C. Prieur. *Basics and Trends in Sensitivity Analysis.* Society for Industrial and Applied Mathematics, 2021

A *pick-freeze* formulation [17] has been proposed to estimate the aggregated Sobol' and β^k indices $\rightarrow$ Large Monte-Carlo samples are needed.

[17] I.M. Sobol. Sensitivity estimates for non linear mathematical models. *Mathematical Modeling and Computer Experiments*, 1:407–414, 1993

[18] L. Le Gratiet. *Multi-fidelity Gaussian process regression for computer experiments.* Phd thesis, Paris 7, January 2013

Kriging-based estimation

A *pick-freeze* formulation [17] has been proposed to estimate the aggregated Sobol′ and β^k indices → Large Monte-Carlo samples are needed.

GP surrogate model uncertainty propagated into the estimates → Sampling GPs on large set of evaluation points.

[17] I.M. Sobol. Sensitivity estimates for non linear mathematical models. *Mathematical Modeling and Computer Experiments*, 1:407–414, 1993

[18] L. Le Gratiet. *Multi-fidelity Gaussian process regression for computer experiments.* Phd thesis, Paris 7, January 2013

Kriging-based estimation

A *pick-freeze* formulation [17] has been proposed to estimate the aggregated Sobol′ and β^k indices → Large Monte-Carlo samples are needed.

GP surrogate model uncertainty propagated into the estimates → Sampling GPs on large set of evaluation points.

↪ Adaptation of the kriging conditioning sampling method detailed in Le Gratiet [2013][18] to noisy observations during the PhD.

[17] I.M. Sobol. Sensitivity estimates for non linear mathematical models. *Mathematical Modeling and Computer Experiments*, 1:407–414, 1993

[18] L. Le Gratiet. *Multi-fidelity Gaussian process regression for computer experiments.* Phd thesis, Paris 7, January 2013

Results on the ASG piping system

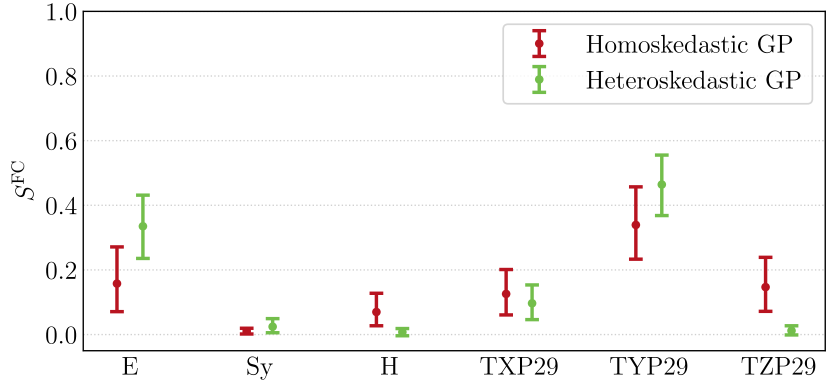
Global sensitivity analysis for 6 mechanical parameters of the structure

Variable name	*Description*
E	Young modulus
Sy	Elasticity limit
H	Hardening module
TPX29	Translation stiffness for the P29 clamped end in X direction
TPY29	Translation stiffness for the P29 clamped end in Y direction
TPZ29	Translation stiffness for the P29 clamped end in Z direction

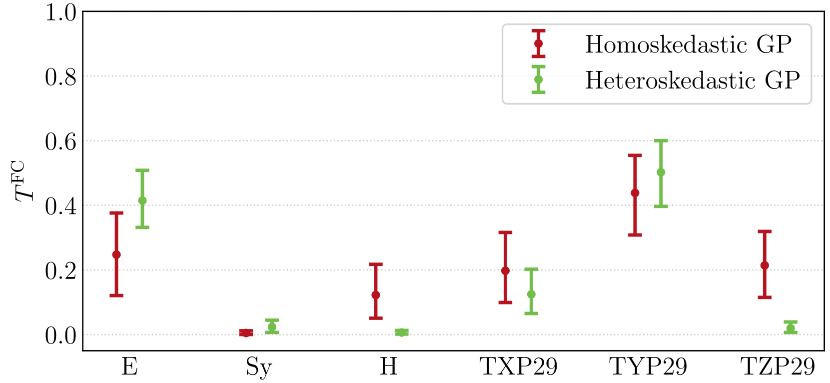
↪ Material parameters

↪ Boundary conditions parameters

Results on the ASG piping system

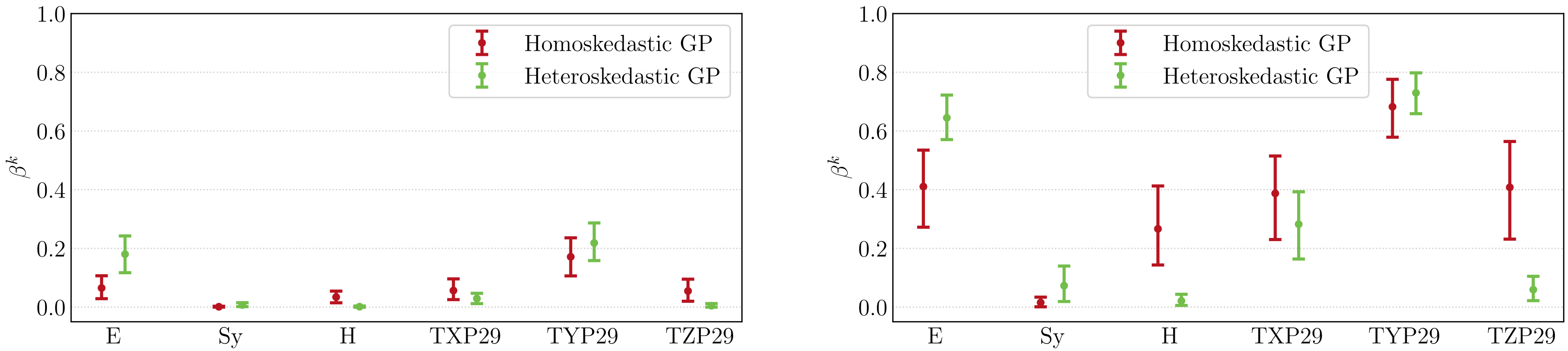


(a) First order



(b) Total order

$\hookrightarrow$ Graphical representation of the posterior distribution of the aggregated Sobol' indices for the ASG piping system using 300 CAST3M computations.



(a) First order (b) Total order

$$k(\Psi, \Psi') = \exp\left(-\frac{\|\Psi - \Psi'\|^2_{L^2_h(\mathcal{A})}}{2\ell^2}\right)$$

↪ Graphical representation of the posterior distribution of the β^k indices for the ASG piping system.

Outline of the defense

Design of experiments

Goal: Improve the estimation accuracy of seismic fragility curves with a fixed budget of evaluations of the computer code $z(a, \mathrm{x})$.

[19] C. Gauchy, C. Feau, and J. Garnier. Importance sampling based active learning for parametric seismic fragility curve estimation.
2021.
doi: 10.48550/ARXIV.2109.04323

[20] J. Bect, D. Ginsbourger, L. Li, V. Picheny, and E. Vazquez. Sequential design of computer experiments for the estimation of a probability of failure.
Statistics and Computing, 22(3):773–793, 2012

Goal: Improve the estimation accuracy of seismic fragility curves with a fixed budget of evaluations of the computer code $\boldsymbol{z}(\boldsymbol{a}, \mathrm{x})$.

$\hookrightarrow$ Development of an importance sampling based active learning algorithm [19].

[19] C. Gauchy, C. Feau, and J. Garnier. Importance sampling based active learning for parametric seismic fragility curve estimation.
2021.
doi: 10.48550/ARXIV.2109.04323

[20] J. Bect, D. Ginsbourger, L. Li, V. Picheny, and E. Vazquez. Sequential design of computer experiments for the estimation of a probability of failure.
Statistics and Computing, 22(3):773–793, 2012

Design of experiments

Goal: Improve the estimation accuracy of seismic fragility curves with a fixed budget of evaluations of the computer code $z(\boldsymbol{a}, \mathrm{x})$.

$\hookrightarrow$ Development of an importance sampling based active learning algorithm [19].

$\hookrightarrow$ Proposition of a Stepwise Uncertainty Reduction (SUR) criterion [20] for seismic fragility curves in the PhD.

[19] C. Gauchy, C. Feau, and J. Garnier. Importance sampling based active learning for parametric seismic fragility curve estimation.
2021.
doi: 10.48550/ARXIV.2109.04323

[20] J. Bect, D. Ginsbourger, L. Li, V. Picheny, and E. Vazquez. Sequential design of computer experiments for the estimation of a probability of failure.
Statistics and Computing, 22(3):773–793, 2012

Given $(\boldsymbol{Y}|\mathcal{D}_n) \sim \mathcal{N}(\widehat{G}_n(a, \mathrm{x}), \widehat{\sigma}_n(a, \mathrm{x})^2)$ and define

$$\widehat{\Psi}_n(a, \mathrm{x}) = \Phi\left(\frac{\widehat{G}_n(a, \mathrm{x}) - \log(C)}{\widehat{\sigma}_n(a, \mathrm{x})}\right) .$$

The sequential procedure is defined by

$$A_{n+1}^{\mathrm{SUR}}, \mathrm{X}_{n+1}^{\mathrm{SUR}} = \underset{a, \mathrm{x} \in \mathcal{A} \times \mathcal{X}}{\operatorname{argmin}} J_n(a, \mathrm{x}) ,$$

where J_n is the SUR sampling criterion at step n defined by

$$J_n(a, \mathrm{x}) = \mathbb{E}_{G}\left[\int_{\mathcal{A} \times \mathcal{X}} (\Psi(\alpha, u; G) - \widehat{\Psi}_{n+1}(\alpha, u))^2 dh(\alpha) d\mathbb{P}_{\mathrm{X}}(u) \middle| A_{n+1} = a, \mathrm{X}_{n+1} = \mathrm{x}, \mathcal{F}_n\right]$$

$\mathcal{F}_n$ is the σ-algebra generated by $\mathcal{D}_n$.

SUR sampling criterion

Given $(\boldsymbol{Y}|\boldsymbol{\mathcal{D}}_n) \sim \boldsymbol{\mathcal{N}}(\widehat{\boldsymbol{G}}_n(\boldsymbol{a},\mathrm{x}), \widehat{\boldsymbol{\sigma}}_n(\boldsymbol{a},\mathrm{x})^2)$ and define

$$\widehat{\boldsymbol{\Psi}}_n(\boldsymbol{a},\mathrm{x}) = \Phi\left(\frac{\widehat{\boldsymbol{G}}_n(\boldsymbol{a},\mathrm{x}) - \log(\boldsymbol{C})}{\widehat{\boldsymbol{\sigma}}_n(\boldsymbol{a},\mathrm{x})}\right) .$$

The sequential procedure is defined by

$$\boldsymbol{A}_{n+1}^{\mathrm{SUR}}, \mathrm{X}_{n+1}^{\mathrm{SUR}} = \underset{\boldsymbol{a},\mathrm{x}\in\boldsymbol{\mathcal{A}}\times\boldsymbol{\mathcal{X}}}{\operatorname{argmin}}\, \boldsymbol{J}_n(\boldsymbol{a},\mathrm{x}) ,$$

where $\boldsymbol{J}_n$ is the SUR sampling criterion at step $\boldsymbol{n}$ defined by

$$\boldsymbol{J}_n(\boldsymbol{a},\mathrm{x}) = \mathbb{E}_{\boldsymbol{G}}\left[\int_{\boldsymbol{\mathcal{A}}\times\boldsymbol{\mathcal{X}}} (\boldsymbol{\Psi}(\boldsymbol{\alpha},\boldsymbol{u};\boldsymbol{G}) - \widehat{\boldsymbol{\Psi}}_{n+1}(\boldsymbol{\alpha},\boldsymbol{u}))^2 d\boldsymbol{h}(\boldsymbol{\alpha}) d\mathbb{P}_{\mathrm{X}}(\boldsymbol{u}) \middle| \boldsymbol{A}_{n+1} = \boldsymbol{a}, \mathrm{X}_{n+1} = \mathrm{x}, \boldsymbol{\mathcal{F}}_n\right]$$

$\boldsymbol{\mathcal{F}}_n$ is the $\boldsymbol{\sigma}$-algebra generated by $\boldsymbol{\mathcal{D}}_n$.

$\hookrightarrow$ Expression of $\boldsymbol{J}_n$ + Practical methods for its computation.

Benchmark settings

- 10 randomly chosen realizations are used for initialization;

Benchmark settings

- 10 randomly chosen realizations are used for initialization;
- At step $\boldsymbol{n}$, $\boldsymbol{m} = \mathbf{1000}$ candidate points $(\boldsymbol{A}_i, \mathbf{X}_i)_{1 \leq i \leq m}$ are subsampled in the dataset of 2000 CAST3M computations. We define:

$$(\boldsymbol{A}_{n+1}^{\mathrm{SUR}}, \mathbf{X}_{n+1}^{\mathrm{SUR}}) = \underset{1 \leq i \leq m}{\operatorname{argmin}} \, \boldsymbol{J}_n(\boldsymbol{A}_i, \mathbf{X}_i) \, .$$

Performance metrics

A numerical benchmark is carried out to compare the performance of SUR strategy and Monte-Carlo designs in terms of posterior variance:

$$\boldsymbol{v_n} = \mathbb{E}_{\boldsymbol{G}}\left[\int_{\boldsymbol{\mathcal{A}}\times\boldsymbol{\mathcal{X}}} (\boldsymbol{\Psi}(\boldsymbol{\alpha}, \boldsymbol{u}; \boldsymbol{G}) - \widehat{\boldsymbol{\Psi}}_n(\boldsymbol{\alpha}, \boldsymbol{u}))^2 \boldsymbol{d}\boldsymbol{h}(\boldsymbol{\alpha})\boldsymbol{d}\mathbb{P}_{\mathrm{X}}(\boldsymbol{u}) \middle| \boldsymbol{\mathcal{F}_n}\right] ,$$

Performance metrics

A numerical benchmark is carried out to compare the performance of SUR strategy and Monte-Carlo designs in terms of posterior variance:

$$\boldsymbol{v_n} = \mathbb{E}_{\boldsymbol{G}}\left[\int_{\boldsymbol{\mathcal{A}}\times\boldsymbol{\mathcal{X}}} (\boldsymbol{\Psi}(\boldsymbol{\alpha}, \boldsymbol{u}; \boldsymbol{G}) - \widehat{\boldsymbol{\Psi}}_{\boldsymbol{n}}(\boldsymbol{\alpha}, \boldsymbol{u}))^2 \boldsymbol{dh}(\boldsymbol{\alpha}) \boldsymbol{d}\mathbb{P}_{\mathrm{X}}(\boldsymbol{u}) \middle| \boldsymbol{\mathcal{F}_n}\right] ,$$

and in terms of bias using a reference fragility curve $\boldsymbol{\Psi}_{\mathrm{ref}}$:

$$\boldsymbol{b_n} = \int_{\boldsymbol{\mathcal{A}}\times\boldsymbol{\mathcal{X}}} (\boldsymbol{\Psi}_{\mathrm{ref}}(\boldsymbol{\alpha}, \boldsymbol{u}) - \widehat{\boldsymbol{\Psi}}_{\boldsymbol{n}}(\boldsymbol{\alpha}, \boldsymbol{u}))^2 \boldsymbol{dh}(\boldsymbol{\alpha}) \boldsymbol{d}\mathbb{P}_{\mathrm{X}}(\boldsymbol{u}) \, .$$

Performance metrics

A numerical benchmark is carried out to compare the performance of SUR strategy and Monte-Carlo designs in terms of posterior variance:

$$\boldsymbol{v_n} = \mathbb{E}_{\boldsymbol{G}}\left[\int_{\mathcal{A}\times\mathcal{X}} (\boldsymbol{\Psi}(\boldsymbol{\alpha}, \boldsymbol{u}; \boldsymbol{G}) - \widehat{\boldsymbol{\Psi}}_{\boldsymbol{n}}(\boldsymbol{\alpha}, \boldsymbol{u}))^2 \boldsymbol{dh}(\boldsymbol{\alpha})\boldsymbol{d}\mathbb{P}_{\mathrm{X}}(\boldsymbol{u}) \middle| \boldsymbol{\mathcal{F}_n}\right] ,$$

and in terms of bias using a reference fragility curve $\boldsymbol{\Psi}_{\mathbf{ref}}$:

$$\boldsymbol{b_n} = \int_{\mathcal{A}\times\mathcal{X}} (\boldsymbol{\Psi}_{\mathbf{ref}}(\boldsymbol{\alpha}, \boldsymbol{u}) - \widehat{\boldsymbol{\Psi}}_{\boldsymbol{n}}(\boldsymbol{\alpha}, \boldsymbol{u}))^2 \boldsymbol{dh}(\boldsymbol{\alpha})\boldsymbol{d}\mathbb{P}_{\mathrm{X}}(\boldsymbol{u}) \,.$$

The integral is evaluated with a Monte-Carlo sample of size 5000 and the expectation on $\boldsymbol{G}$ using 4000 realizations of the GP surrogate.

Benchmark setting

The SUR strategy is compared to a Monte-Carlo design.

Benchmark setting

The SUR strategy is compared to a Monte-Carlo design.

100 replications of Monte-Carlo designs for several training sizes are computed.

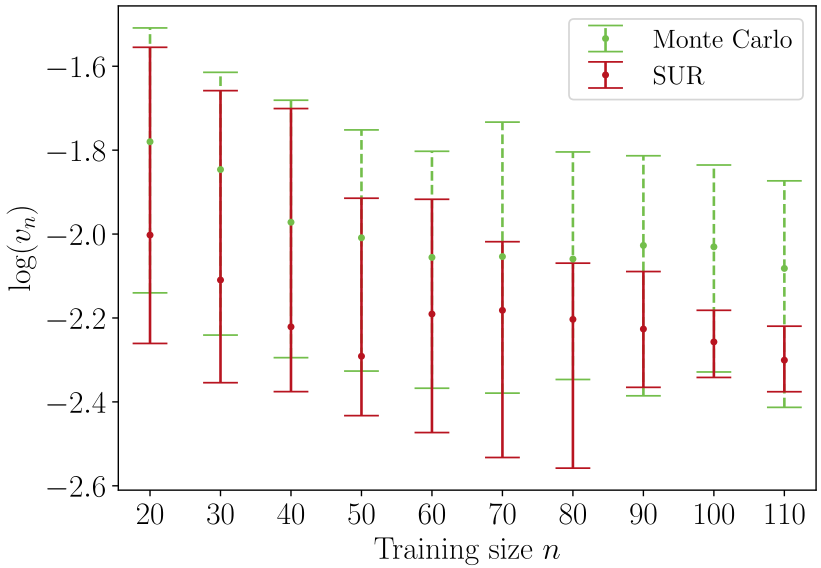
Benchmark setting

The SUR strategy is compared to a Monte-Carlo design.

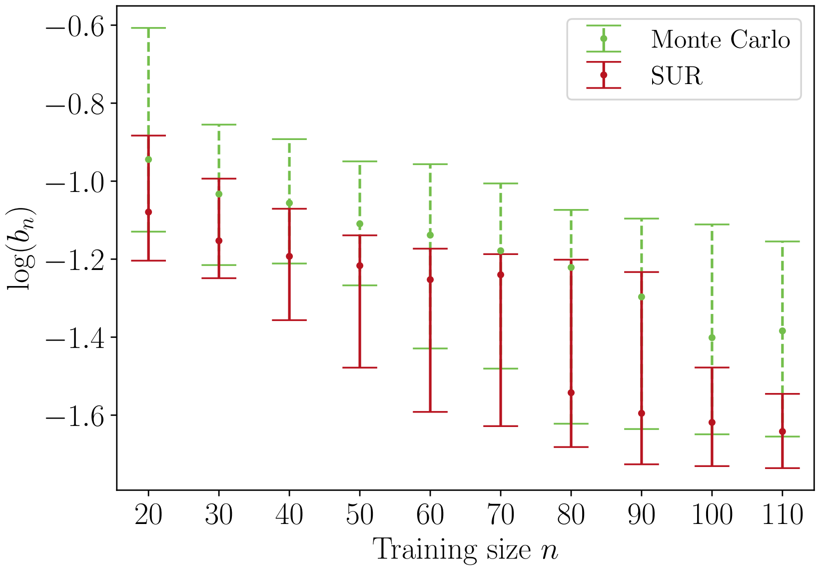
100 replications of Monte-Carlo designs for several training sizes are computed.

Due to the randomness induced in the SUR algorithm by choosing the candidate points at each step, 100 runs of the SUR strategy are carried out using HPC.

Performance assessment



(a) Performance evaluation on the variance



(b) Performance evaluation on the bias

1. Uncertainty Quantification
2. Surrogate modeling of computer codes using Gaussian processes
3. Global sensitivity analysis on seismic fragility curves
4. Sequential design of experiments
5. Conclusion and perspectives

Contributions

- Uncertainty Quantification framework development for earthquake engineering;
- Estimators of seismic fragility curve based on GP surrogates;
- Global sensitivity indices defined on seismic fragility curves;
- Sequential design of experiments
  ↪ Importance sampling based active learning (frequentist viewpoint)

  ↪ SUR procedure (Bayesian viewpoint)

C. Gauchy, C. Feau, and J. Garnier. Importance sampling based active learning for parametric seismic fragility curve estimation.
2021.
doi: 10.48550/ARXIV.2109.04323

C. Gauchy, C. Feau, and J. Garnier. Uncertainty quantification and global sensitivity analysis of seismic fragility curves using kriging.
2022a.
doi: 10.48550/ARXIV.2210.06266

C. Gauchy, C. Feau, and J. Garnier. Estimation of seismic fragility curves by sequential design of experiments.
In *52ème journées de Statistique de la Société Française de Statistique (JdS 2022)*, Lyon, France, 2022b

Perspectives

- Latent model for heteroskedastic GP regression [21];
- Model selection (AIC, BIC, Bayes factor,...);
- Bayesian methodology for seismic fragility curve estimation. Prior elicitation using reference prior theory [22].

[21] A. Marrel, B. Iooss, S. Da Veiga, and M. Ribatet. Global sensitivity analysis of stochastic computer models with joint metamodels.
Statistics and Computing, 22(3):833–847, 2012

[22] J. O. Berger, J. M. Bernardo, and D. Sun. The formal definition of reference priors.
The Annals of statistics, 37(2):905–938, 05 2009

Thank you for your attention !

`clgch.github.io`

Bibliography I

J. Barr and H. Rabitz. A generalized kernel method for global sensitivity analysis. *SIAM/ASA Journal on Uncertainty Quantification*, 10(1):27–54, 2022.

J. Bect, D. Ginsbourger, L. Li, V. Picheny, and E. Vazquez. Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing*, 22(3):773–793, 2012.

J. O. Berger, J. M. Bernardo, and D. Sun. The formal definition of reference priors. *The Annals of statistics*, 37(2):905–938, 05 2009.

S. Da Veiga. Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, 85(7):1283–1305, 2015.

S. Da Veiga, F. Gamboa, B. Iooss, and C. Prieur. *Basics and Trends in Sensitivity Analysis*. Society for Industrial and Applied Mathematics, 2021.

E. De Rocquigny, N. Devictor, S. Tarantola, Y. Lefebvre, N. Pérot, W. Castaings, F. Mangeant, C. Schwob, R. Lavin, J.R. Masse, P. Limbourg, W. Kanning, and P.H.A.J.M. Gelder. *Uncertainty in industrial practice: A guide to quantitative uncertainty management*. Wiley, 2008.

A. Der Kiureghian and O. Ditlevsen. Aleatory or epistemic? does it matter? *Structural Safety*, 31(2):105–112, 2009.

Bibliography II

C. Gauchy, C. Feau, and J. Garnier. Importance sampling based active learning for parametric seismic fragility curve estimation. 2021. doi: 10.48550/ARXIV.2109.04323.

C. Gauchy, C. Feau, and J. Garnier. Uncertainty quantification and global sensitivity analysis of seismic fragility curves using kriging. 2022a. doi: 10.48550/ARXIV.2210.06266.

C. Gauchy, C. Feau, and J. Garnier. Estimation of seismic fragility curves by sequential design of experiments. In *52ème journées de Statistique de la Société Française de Statistique (JdS 2022)*, Lyon, France, 2022b.

A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola. A kernel two-sample test. *Journal of Machine Learning Research*, 13(25):723–773, 2012.

P. Hall, C.C. Heyde, Z.W. Birnbaum, and E. Lukacs. *Martingale Limit Theory and Its Application*. Communication and Behavior. Elsevier Science, 2014.

B. Iooss and L. Le Gratiet. Uncertainty and sensitivity analysis of functional risk curves based on Gaussian processes. *Reliability Engineering & System Safety*, 187: 58–66, 2019.

Robert P. Kennedy. Risk based seismic design criteria. *Nuclear Engineering and Design*, 1999.

R.P. Kennedy, C.A. Cornell, R.D. Campbell, S. Kaplan, and H.F. Perla. Probabilistic seismic safety study of an existing nuclear power plant. *Nuclear Engineering and Design*, 59(2):315 – 338, 1980.

A. P. Kyprioti and A. A. Taflanidis. Kriging metamodeling for seismic response distribution estimation. *Earthquake Engineering & Structural Dynamics*, 50(13):3550–3576, 2021.

L. Le Gratiet. *Multi-fidelity Gaussian process regression for computer experiments*. Phd thesis, Paris 7, January 2013.

A. Marrel, B. Iooss, S. Da Veiga, and M. Ribatet. Global sensitivity analysis of stochastic computer models with joint metamodels. *Statistics and Computing*, 22(3):833–847, 2012.

A. Marrel, B. Iooss, and V. Chabridon. The icscream methodology: Identification of penalizing configurations in computer experiments using screening and metamodel—applications in thermal-hydraulics. *Nuclear Science and Engineering*, 196(3): 301–321, 2022.

S. Rezaeian and A. Der Kiureghian. Simulation of synthetic ground motions for specified earthquake and site characteristics. *Earthquake Engineering & Structural Dynamics*, 39(10):1155–1180, 2010.

Bibliography IV

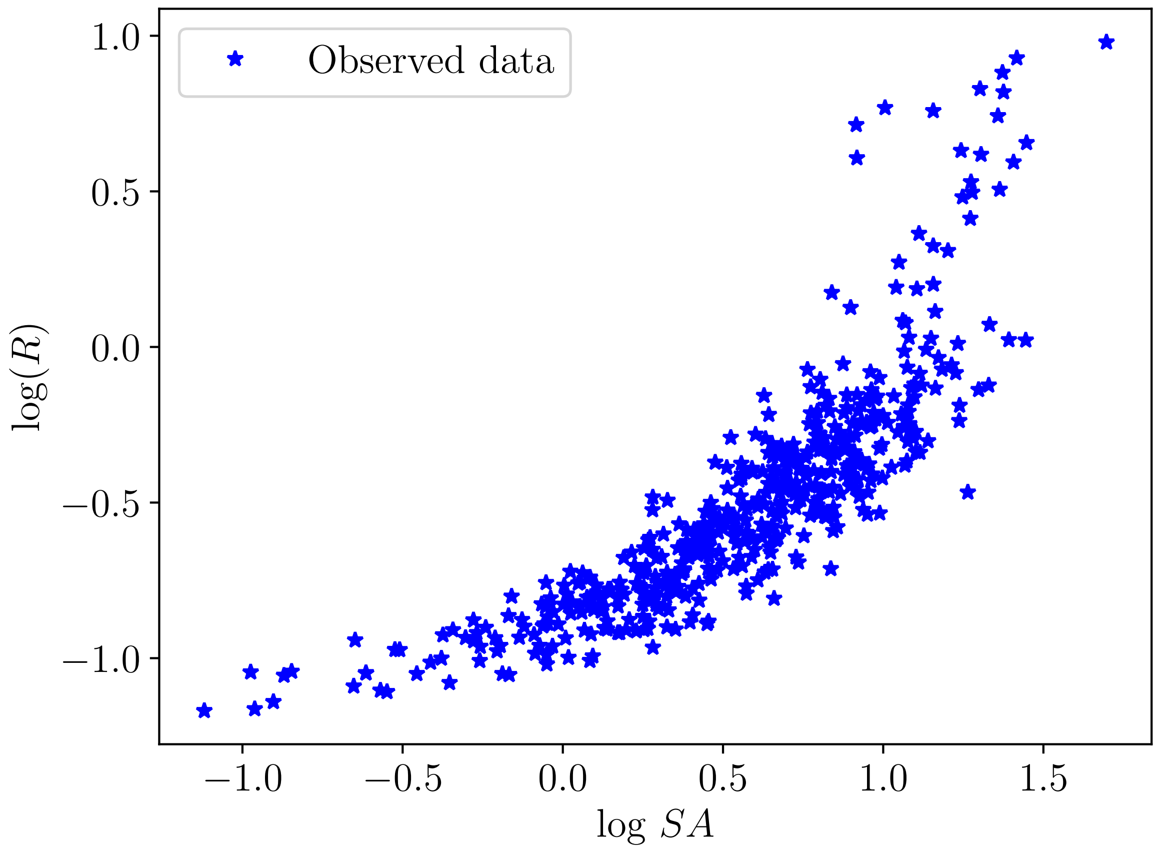
I.M. Sobol. Sensitivity estimates for non linear mathematical models. *Mathematical Modeling and Computer Experiments*, 1:407–414, 1993.

C. Soize and R. Ghanem. Physical systems with random uncertainties: Chaos representations with arbitrary probability measure. *SIAM Journal on Scientific Computing*, 26(2):395–410, 2004.

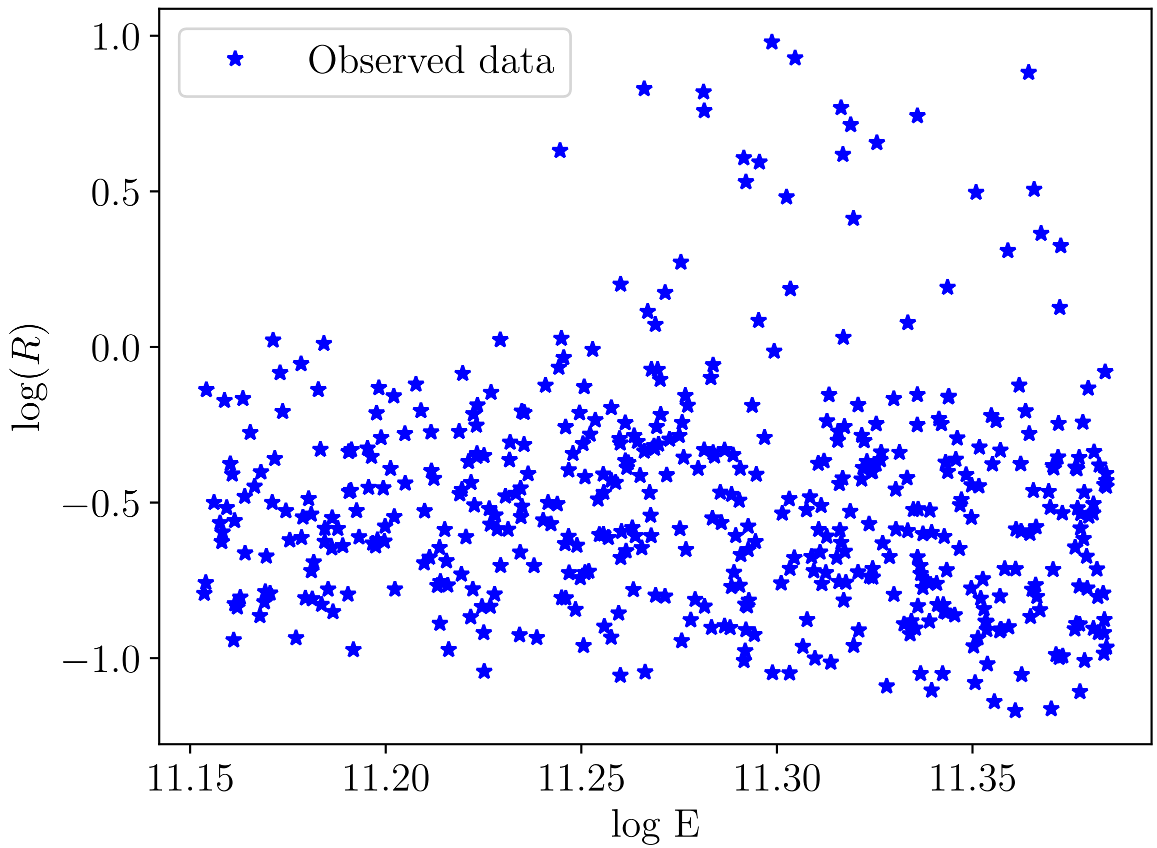
F. Touboul, P. Sollogoub, and N. Blay. Seismic behaviour of piping systems with and without defects: experimental and numerical evaluations. *Nuclear Engineering and Design*, 192(2):243–260, 1999.

A. W. Van der Vaart. *Asymptotic Statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998.

ASG data

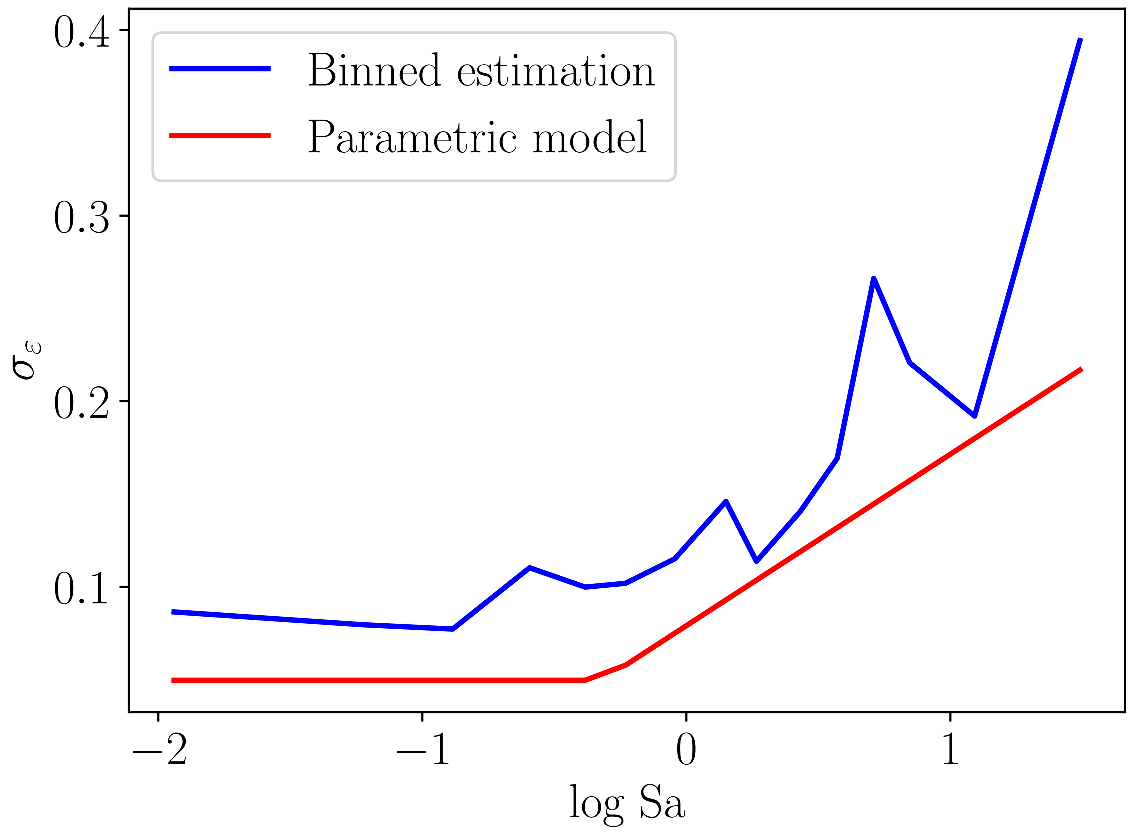


ASG data

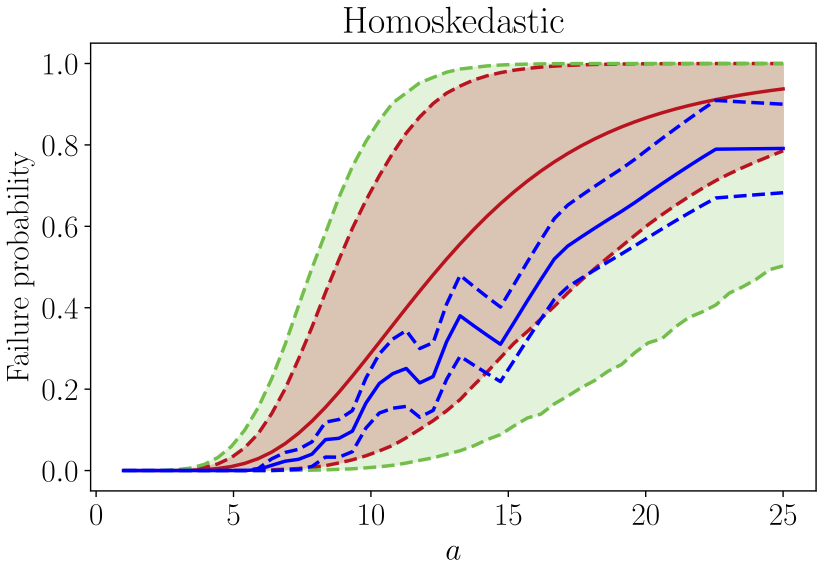


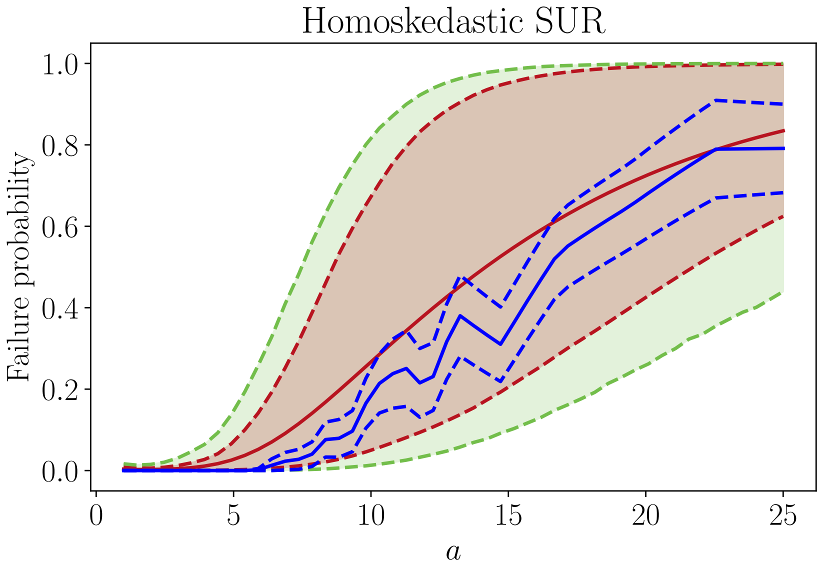
Observed data
1.0
0.5
0.0
−0.5
−1.0
log(R)
11.15
11.20
11.25
11.30
11.35
log E

Heteroskedastic noise



SUR comparison





Parametric estimation

Given data $(\boldsymbol{X}_i, \boldsymbol{S}_i)_{1\leq i\leq n}$ where $\boldsymbol{S}_i \in \{\boldsymbol{0}, \boldsymbol{1}\}$ and $\boldsymbol{X}_i = \log(\boldsymbol{A}_i)$. The fragility curve $\boldsymbol{\psi}(\boldsymbol{x})$ is just

$$\boldsymbol{\psi}(\boldsymbol{x}) = \mathbb{E}[\boldsymbol{S}|\boldsymbol{X} = \boldsymbol{x}]\,.$$

Given a function space $\boldsymbol{\mathcal{F}} = \{\boldsymbol{\psi}_{\boldsymbol{\theta}},\ \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$. The goal is to minimize

$$\boldsymbol{g}(\boldsymbol{\theta}) = \mathbb{E}[(\boldsymbol{\psi}(\boldsymbol{X}) - \boldsymbol{\psi}_{\boldsymbol{\theta}}(\boldsymbol{X}))^2]\,.$$

In practice optimization is done on the following objective function

$$\boldsymbol{r}(\boldsymbol{\theta}) = \mathbb{E}[(\boldsymbol{S} - \boldsymbol{\psi}_{\boldsymbol{\theta}}(\boldsymbol{X}))^2]\,.$$

Empirical estimator:

$$\widehat{\boldsymbol{R}}_{\boldsymbol{n}}(\boldsymbol{\theta}) = \frac{\boldsymbol{1}}{\boldsymbol{n}}\sum_{\boldsymbol{i}=\boldsymbol{1}}^{\boldsymbol{n}} (\boldsymbol{S}_i - \boldsymbol{\psi}_{\boldsymbol{\theta}}(\boldsymbol{X}_i))^2\ .$$

IS-AL principle

$\hookrightarrow$ Importance sampling to reduce the variance of the empirical approximation of $r(\theta)$.

$$\widehat{R}_n^{\mathrm{IS}}(\theta) = \frac{1}{n}\sum_{i=1}^{n} \frac{p(X_i)}{q(X_i)} \left(S_i - \psi_\theta(X_i)\right)^2 \ .$$

The instrumental density q is chosen to minimize the variance:

$$\int_{\mathcal{X}} \frac{p^2(x)}{q(x)} \tilde{\ell}_\theta^2(x) dx - r(\theta)^2 \ ,$$

where $\tilde{\ell}_\theta^2(x) = \mathbb{E}[(S - \psi_\theta(X))^4 | X = x]$.

Parametric family of lognormal seismic fragility curves: $\mathcal{F} = \{\Phi\left(\frac{\log(IM/\alpha)}{\beta}\right), (\alpha, \beta) \in \Theta\}$

Penalization term $\Omega(\theta) = \frac{\beta_{\text{reg}}}{\beta}$

$\hookrightarrow$ Importance sampling to reduce the variance of the empirical approximation of $r(\theta)$.

$$\widehat{R}_n^{\text{IA}}(\theta) = \frac{1}{n} \sum_{i=1}^{n} \frac{p(X_i)}{q_{\widehat{\theta}_{i-1}^{\text{IA}}, \varepsilon}(X_i)} \left(S_i - \psi_\theta(X_i)\right)^2 + \frac{\beta_{\text{reg}}}{n\beta} ,$$

$$\widehat{\theta}_n^{\text{IA}} = \underset{\theta \in \Theta}{\operatorname{argmin}} \, \widehat{R}_n^{\text{IA}}(\theta)$$

Theoretical results

$\hookrightarrow$ Consistency of $\widehat{\boldsymbol{\theta}}_n^{\mathrm{IA}}$. Let $\boldsymbol{\theta}_* = \operatorname{argmin} \mathbb{E}[(\psi(\boldsymbol{X}) - \psi_{\boldsymbol{\theta}}(\boldsymbol{X}))^2]$,

$$\widehat{\boldsymbol{\theta}}_n^{\mathrm{IA}} \xrightarrow[n \to +\infty]{} \boldsymbol{\theta}_* \text{ in probability.}$$

$\hookrightarrow$ Asymptotic normality of $\widehat{\boldsymbol{\theta}}_n^{\mathrm{IA}}$.

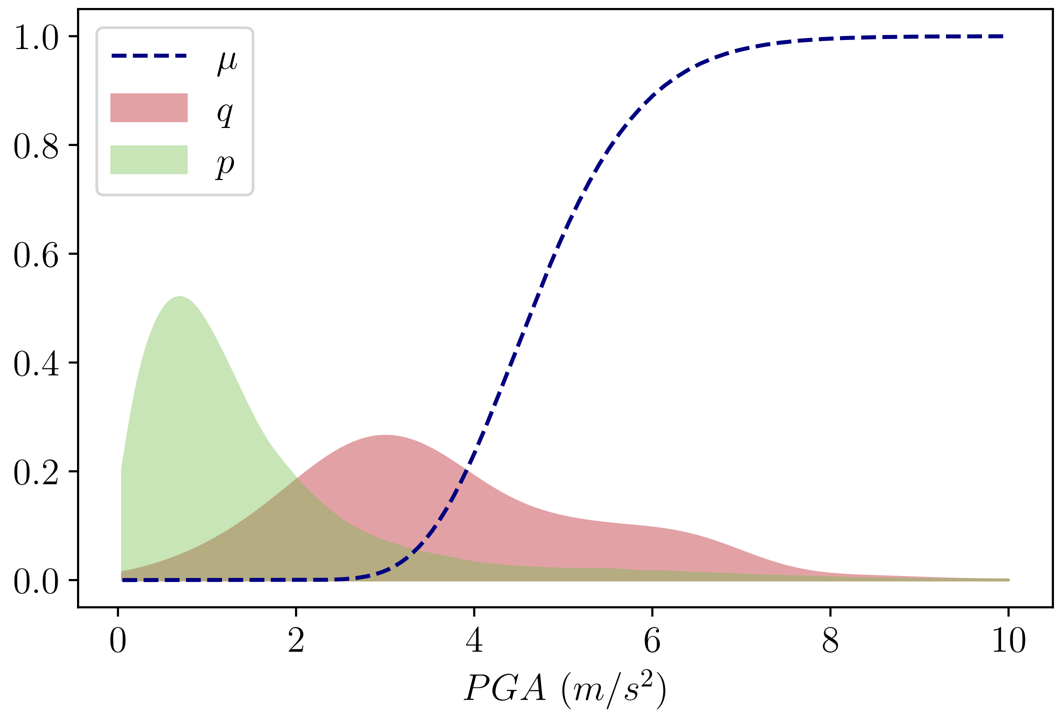
$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_n^{\mathrm{IA}} - \boldsymbol{\theta}_*) \xrightarrow{\mathcal{D}} \mathcal{N}(0, G_{\boldsymbol{\theta}_*,\varepsilon}^{-1})$$

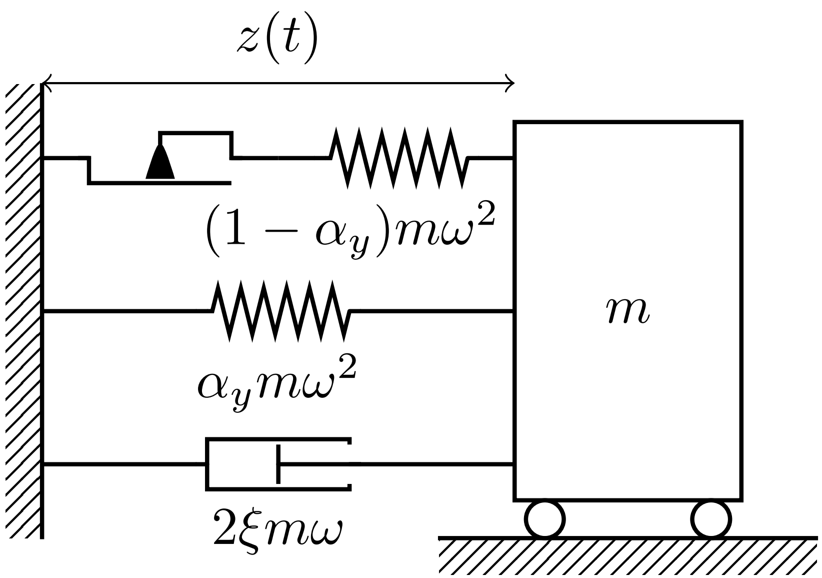
Proofs sketch: Combine convergence of martingal results (Hall et al. [2014][23]) with Z-estimation theory (Van der Vaart [1998][24])

[23] P. Hall, C.C. Heyde, Z.W. Birnbaum, and E. Lukacs. *Martingale Limit Theory and Its Application.* Communication and Behavior. Elsevier Science, 2014

[24] A. W. Van der Vaart. *Asymptotic Statistics.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 1998

IS-AL interpretation





$$\ddot{\boldsymbol{z}}(\boldsymbol{t}) + \boldsymbol{2\xi\omega}\dot{\boldsymbol{z}}(\boldsymbol{t}) + \boldsymbol{f}^{\mathrm{NL}}(\boldsymbol{z}(\boldsymbol{t})) = -\boldsymbol{s}(\boldsymbol{t}) \ ,$$

where $\boldsymbol{f}^{\mathrm{NL}}$ is a nonlinear restoring force.

MLE versus IS-AL