DE LA RECHERCHE À L'INDUSTRIE

Polynomial chaos-based uncertainty quantification of material parameters

May 14, 2022

Loïc GIRALDI – IRESNE / DEC / SESC / LSC

Loic.GIRALDI@cea.fr

Short bio

- **CV**
  - **2009-2015 @ Centrale Nantes : PhD + post-doc**
    - Uncertainty propagation for high dimensional problems
    - Semi-intrusive uncertainty propagation
  - **2015-2018 @ KAUST : Post-doc**
    - Uncertainty quantification and data assimilation applied to geosciences
  - **2018-2020 @ Renault : Data Scientist**
    - Uncertainty quantification and AI for the digital validation of autonomous vehicles
  - **Since 2020/12/21 @ CEA : Research Scientist**
    - Uncertainty quantification and AI for nuclear fuel performance simulations

- **Research topics**
  - Uncertainty quantification
  - AI and data assimilation
  - Numerical methods for model order reduction

Uncertainty quantification

- **Uncertainty quantification definition**
  - Science of **quantifying** and **reducing** uncertainties in computational and real-world systems

- **Classification of uncertainty sources**
  - **Epistemic uncertainty:** uncertainty that comes from a lack of knowledge, **can be reduced**
    - **Example:** my height, material parameters,...
  - **Aleatoric uncertainty:** uncertainty that differs each time we run the experiment, **cannot be reduced**
    - **Example:** rolling a dice,...

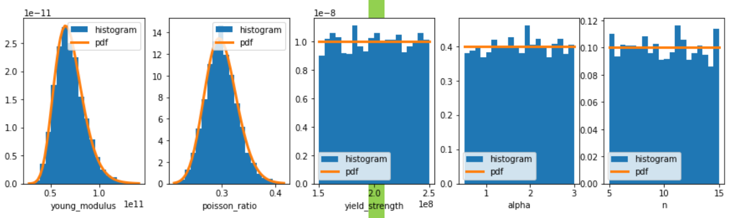
- **Aim of the course**
  - Perform the uncertainty quantification of material parameters from a tensile test and a shear test
    - Knowledge of the uncertainty quantification process
    - Perform the different steps of the above process

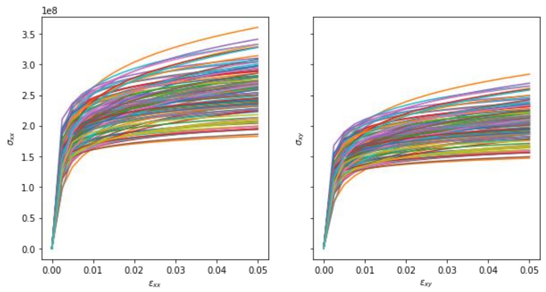
Uncertainty quantification process overview

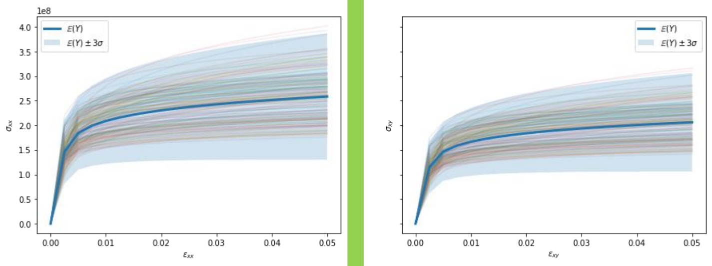
STEP B
Uncertainty quantification
Modeling uncertainty sources with probability distribution

STEP A
Problem specification
Computational model definition
Assessment criteria definition
$Y = M(X)$

STEP C
Uncertainty propagation







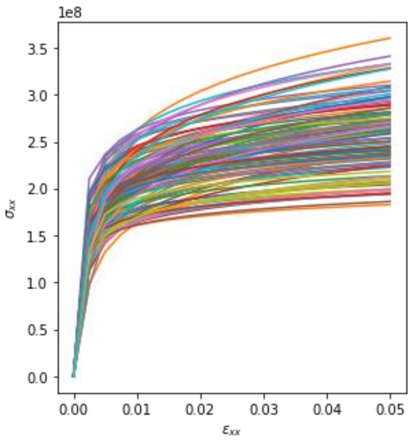
STEP C'
Sensitivity analysis
Ranking of the sources of uncertainty

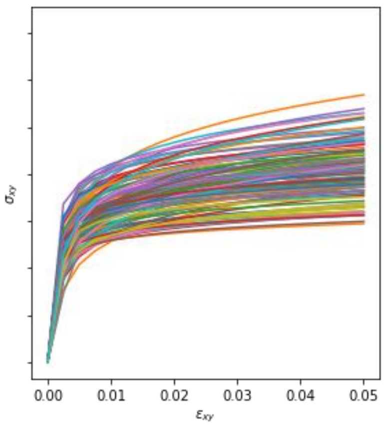
STEP B'
Quantification of sources
Inverse methods, Bayesian inference

STEP A: Problem specification

- Simulation of a tensile test and a shear test with MFront / Mtest (tfel.sourceforge.net/index.html)
- Nonlinear elastic Ramberg-Osgood constitutive behavior

$$\varepsilon = \frac{1}{3K} tr(\sigma) I + \frac{\sigma_{VM}}{3\mu} n_{VM} + \alpha \left(\frac{\sigma_{VM}}{\sigma_0}\right)^n n_{VM}.$$





- where
  - E is the Young modulus,
  - ν is the Poisson ratio,
  - $K = E/(3(1-2\nu))$ is the bulk modulus,
  - $\mu = E/(2(1+\nu))$ is the shear modulus,
  - σ_0 is the yield strength, α and n are numerical parameters describing the pseudo-plastic part of the behavior,
  - σ_{VM} is the Von Mises equivalent stress,
  - n_{VM} is the Von Mises normal

- Tensile test : displacement / strains are imposed such that $\varepsilon = \begin{pmatrix} \varepsilon_{xx} & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{pmatrix}$, and we measure σ_{xx}

- Shear test : displacement / strains are imposed such that $\varepsilon = \begin{pmatrix} * & \varepsilon_{xy} & 0 \\ \varepsilon_{xy} & * & 0 \\ 0 & 0 & * \end{pmatrix}$, and we measure σ_{xy}

- Computational model

$$X = (E, \nu, \sigma_0, \alpha, n) \mapsto Y(X) = \left(\sigma_{xx}^0, \dots, \sigma_{xx}^t, \sigma_{xy}^0, \dots, \sigma_{xy}^t\right) \in \mathbb{R}^T, \qquad T = 2(t+1)$$

Material parameters

- $\boldsymbol{E}$ is the Young modulus,
- $\boldsymbol{\nu}$ is the Poisson ratio,
- $\boldsymbol{\sigma_0}$ is the yield strength,
- $\boldsymbol{\alpha}$ and $\boldsymbol{n}$ are numerical parameters describing the pseudo-plastic part of the behavior.

Probability distribution of the parameters

- The Young modulus $\boldsymbol{E}$ is drawn according to **a log-normal distribution**,
- The Poisson ratio $\boldsymbol{\nu}$ is drawn according to **a log-normal distribution**,
- The parameters $\boldsymbol{\sigma_0, \alpha, n}$ are assumed to be **uniform.**

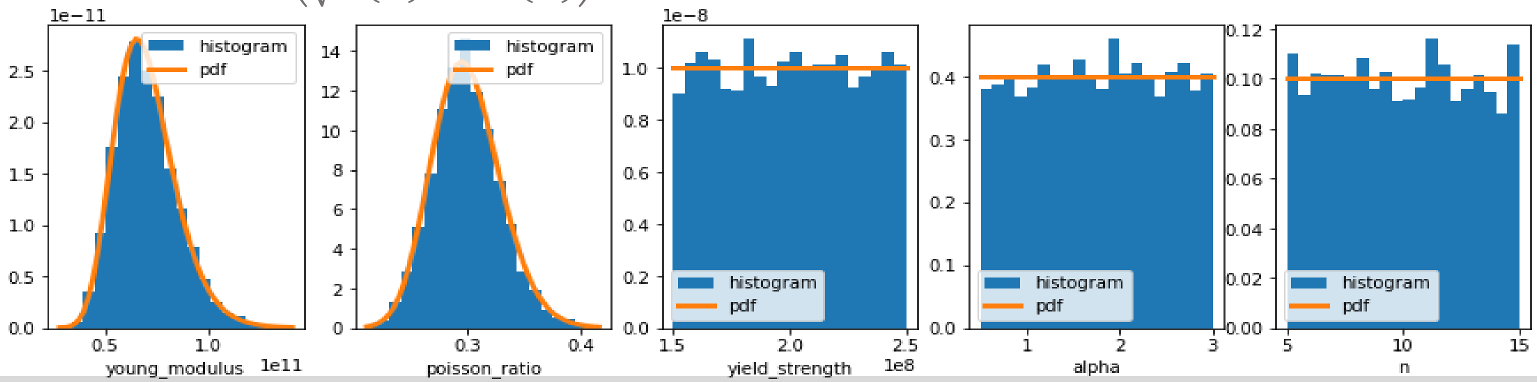
Log-normal distribution crash course

- X follows a log-normal distribution if

$$X = \exp(\mu + \sigma Z) \qquad \textit{where} \qquad Z \sim N(0,1)$$

- We can fully describe a log normal distribution through its mean $\mathbb{E}(X)$ and variance $\mathbb{V}(X)$ with the relations

$$\mu = \log\left(\frac{\mathbb{E}(X)^2}{\sqrt{\mathbb{E}(X)^2 + \mathbb{V}(X)}}\right), \qquad \textit{and} \qquad \sigma = \log\left(1 + \frac{\mathbb{V}(X)}{\mathbb{E}(X)^2}\right).$$



STEP C: Polynomial chaos expansion

- **Issues**
  - Evaluating statistics, probabilities, quantities of interest requires **a large number of evaluations** of the model
  - Optimization based approaches require the **gradient** of the model

- **Solution**
  - **Surrogate modeling (metamodeling)** in order to replace the computational model
  - Approaches in the literature : **polynomial chaos expansion**, Gaussian processes, **machine learning models**, ...

- **Generalized Polynomial Chaos Expansion**
  - Assume that the random variable Y has a finite variance, then Y admits the decomposition

$$Y(X) = \sum_{k=0}^{\infty} w_k \phi_k(X).$$

  - $(\phi_k)_{k=0}^{\infty}$ is a family of multivariate orthogonal polynomials for the L^2 inner product

$$\langle \phi_k, \phi_l \rangle = \mathbb{E}(\phi_k \phi_l) = \int_{\mathcal{X}} \phi_k(x)\phi_l(x)p(x)dx = \delta_{kl}\|\phi_k\|^2.$$

STEP C: Truncation of the polynomial chaos expansion

- **Expression of the truncated polynomial chaos expansion for independent random variables**
  - $X_1, \dots, X_d$ are assumed to be independent random variables such that the basis becomes
  $$\phi_\alpha(X) = \phi^1_{\alpha_1}(X_1) \dots \phi^d_{\alpha_d}(X_d)$$
  - where $\left(\phi^i_k\right)_{k=0}^{\infty}$ is family of univariate polynomials orthogonal with respect to the density $p(X_i)$
  - We assume that $\deg\left(\phi^i_k\right) = k$
  - We then discretize the series such that
  $$Y(X) \approx Y_{PC}(X) = \sum_{k=0}^{P} w_k \phi_k(X) = \sum_{\alpha \in \mathcal{J}} w_\alpha \phi^1_{\alpha_1}(X_1) \dots \phi^d_{\alpha_d}(X_d)\,.$$
  - Where $\mathcal{J} \subset \mathbb{N}^d$ is a **finite set** of multi-indices such that $\#\mathcal{J} = P + 1$.
- **Some truncation strategies – definition of the multi-index set**
  - Full tensorization of the polynomial basis
  $$\mathcal{J} = \{\alpha \in \mathbb{N}^d; \alpha_i \leq D_i\}$$
  - Total degree based truncation
  $$\mathcal{J} = \{\alpha \in \mathbb{N}^d; \sum \alpha_i \ \leq D\}$$
- **Choice of the univariate polynomials**

Probability distribution	Orthogonal polynomials
Gaussian distribution $\mathcal{N}(0,1)$	Hermite
Uniform distribution $\mathcal{U}(-1,1)$	Legendre

STEP C: Polynomial chaos computation in practice

1. **Transform the random variable / dataset**
   - For log-normal distributions $X = \exp(\mu + \sigma Z)$, we have $Z = \frac{\log(X)-\mu}{\sigma} \sim \mathcal{N}(0,1)$
   - For uniform distributions $X = \mathcal{U}(a,b)$, we have $Z = \frac{b-a}{2}X + \frac{b+a}{2} \sim \mathcal{U}(-1,1)$
   - Now, we look for an expansion of the form

$$Y(Z) \approx Y_{PC}(Z) = \sum_{\alpha \in \mathcal{J}} w_\alpha \phi_\alpha(Z)$$

2. **Pick the right polynomials associated to the reference distributions**
3. **Define the truncation scheme via the set of multi-indices $\mathcal{J}$**
4. **Compute the optimal weights $(w_\alpha)_{\alpha \in \mathcal{J}}$ by solving the regression problem given a dataset $(z_i, y_i)_{i=1}^{N}$**

$$\min_w \frac{1}{N} \sum_{i=0}^{N} \left\| y_i - \sum_{\alpha \in \mathcal{J}} w_\alpha \phi_\alpha(z_i) \right\|^2$$

▶ **Model selection is performed via hold-out validation**

Dataset splitting:

Training set	Validation set	Test set
Computation of w	Selection of the best model (degree, truncation,...)	Assessment of the quality of the selected model

STEP C: Polynomial chaos post-processing

- **Given a polynomial chaos expansion**

$$Y_{PC}(Z) = \sum_{\alpha \in \mathcal{J}} w_\alpha \phi_\alpha(Z)$$

- **We can compute some interesting quantities of interest by post-processing**

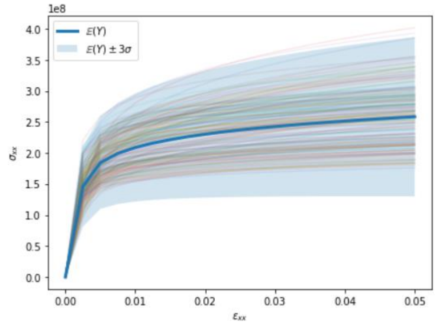
- **Evaluating the model at a sample z requires the evaluation of the basis $\left(\phi_\alpha(z)\right)_{\alpha \in \mathcal{J}}$ and computing the product with the weights $(w_\alpha)_{\alpha \in \mathcal{J}}$**

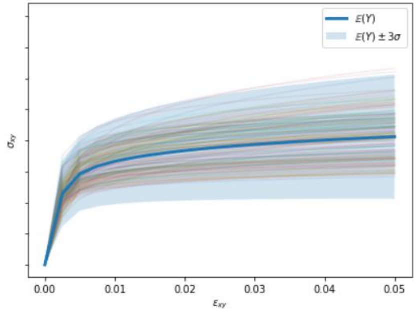
- **The expection is**

$$\mathbb{E}\left(Y_{PC}(Z)\right) = w_0$$

- **The variance is**

$$\mathbb{V}\left(Y_{PC}(Z)\right) = \sum_{\alpha \in \mathcal{J}\setminus\{0\}} w_\alpha^2 \|\phi_\alpha\|^2$$





- **The gradient of the polynomial chaos**

$$\frac{\partial Y_{PC}}{\partial Z_i}(Z) = \sum_{\alpha \in \mathcal{J}} w_\alpha \phi^1_{\alpha_1}(Z_1) \ldots \left(\boldsymbol{\phi^i_{\alpha_i}}\right)'(Z_i) \ldots \phi^d_{\alpha_d}(Z_d)$$

STEP C’: Global sensitivity analysis

- **Variance-based sensitivity analysis:** quantify the proportions of the model output variance induced by the uncertainties of each input parameter
- **The Sobol decomposition exists and is unique for random variables with finite variance and orthogonality condition**

$$Y(Z) = Y_0 + \sum_{i=1}^{d} Y_i(Z_i) + \sum_{i<j}^{d} Y_{ij}\left(Z_i, Z_j\right) + \cdots + Y_{1\ldots d}(Z_1, \ldots Z_d)$$

- **The orthogonality condition implies that we can decompose the variance such that**

$$\mathbb{V}(Y) = \sum_{i=1}^{d} \mathbb{V}(Y_i) + \sum_{i<j}^{d} \mathbb{V}(Y_{ij}) + \cdots + \mathbb{V}(Y_{1\ldots d})$$

$$\text{or,} \qquad 1 = \sum_{i=1}^{d} S_i + \sum_{i<j}^{d} S_{ij} + \cdots + S_{1\ldots d} \text{ where } S_u = \frac{\mathbb{V}(Y_u)}{\mathbb{V}(Y)}$$

- **The quantities $(S_u)_u$ are called first-order Sobol indices and quantify the proportion of the variance of Y induced by the variables in u.**
- **S_i quantify the variance induced by the variable Z_i alone.**
- **The quantity of variance induced by Z_i and its interactions with the other variables is the total order Sobol index T_i defined by**

$$T_i = \sum_{\substack{u \\ i \in u}} S_u$$

STEP C': Sobol indices and polynomial chaos expansion

- **Sobol indices are post-processes from polynomial chaos expansions**

$$Y_{PC}(Z) = \sum_{\alpha \in \mathcal{J}} w_\alpha \phi^1_{\alpha_1}(Z_1) \dots \phi^d_{\alpha_d}(Z_d)$$

- **The variance is**

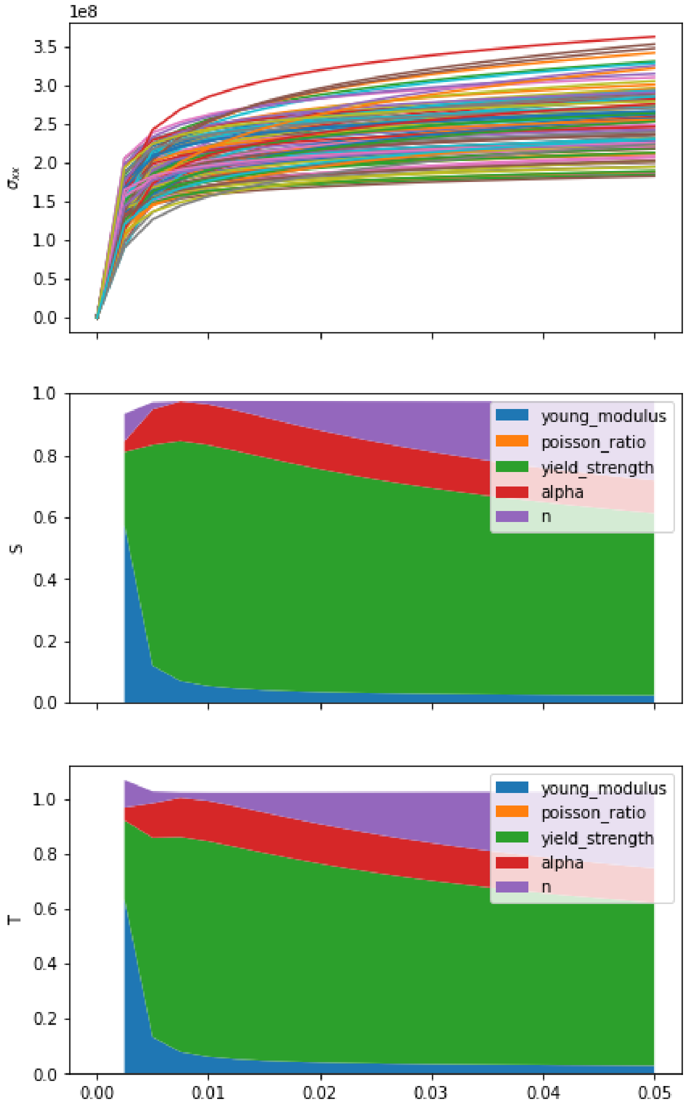
$$\mathbb{V}\big(\boldsymbol{Y_{PC}}(\boldsymbol{Z})\big) = \sum_{\alpha \in \mathcal{J}\setminus\{0\}} w_\alpha^2 \|\phi_\alpha\|^2$$

- **The first-order Sobol index S_i is**

$$S_i = \frac{1}{\mathbb{V}\big(Y_{PC}(Z)\big)} \sum_{\substack{\alpha \in \mathcal{J} \\ \alpha=(0,\dots,0,\alpha_i,0,\dots,0)}} w_\alpha^2 \|\phi_\alpha\|^2$$

- **The total-order Sobol index T_i is**

$$T_i = \frac{1}{\mathbb{V}\big(Y_{PC}(Z)\big)} \sum_{\substack{\alpha \in \mathcal{J} \\ \alpha_i > 0}} w_\alpha^2 \|\phi_\alpha\|^2$$



- **Issue**
  - The distribution of X (or Z) was defined at the beginning.
  - An experiment was performed and some data y_e were collected. The experimental process is assumed to satisfy

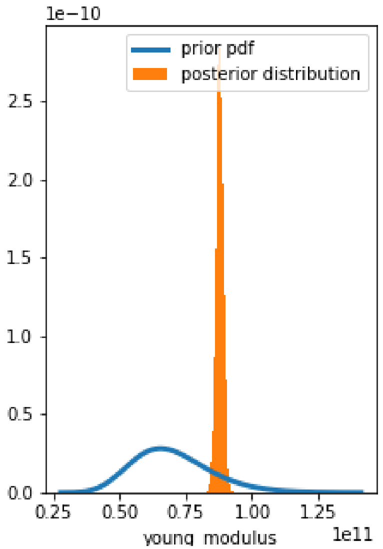
$$Y_e = Y(Z) + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0, \sigma^2 \mathbb{I})$$

  - How can we **update the distribution** of the parameter X **after having observed the data**?

- **Bayes theorem states that**

$$p(Z \mid Y_e = y_e) = \frac{p(Y_e = y_e \mid Z)p(Z)}{p(Y_e = y_e)}$$

  - $p(Z \mid Y_e = y_e)$ is the **posterior distribution**, i.e. the updated distribution of Z after seeing the data y_e,
  - $p(Y_e = y_e \mid Z)$ is the **likelihood**, describing the probability of the data $Y_e = y_e$ knowing the parameter Z
  - $p(Z)$ is the **prior distribution**, the distribution of Z before having observed any data
  - $p(Y_e = y_e)$ is the **model evidence**, the probability of the data (scaling factor).



- **Bayesian inference is a method to update the distribution of Z after knowing some data.**

- **Common objectives related to Bayesian inference**
  - Estimate the value z_{MAP} with the maximum posterior distribution density
  - Approximate the posterior distribution
  - ~~Sample from the posterior distribution (MCMC)~~

STEP B’: MAP estimate

- **Objective**
  - Compute the **Maximum A Posteriori estimate** z_{MAP} maximizing the posterior distribution density such that

$$\log p(\, z_{MAP} \mid Y_e = y_e \,) = \max_{z} \log p(\, z \mid Y_e = y_e \,)$$

- **Since $Y_e = Y(Z) + \varepsilon$ with $\varepsilon \sim \mathcal{N}(0, \sigma^2 \mathbb{I})$, the likelihood is**

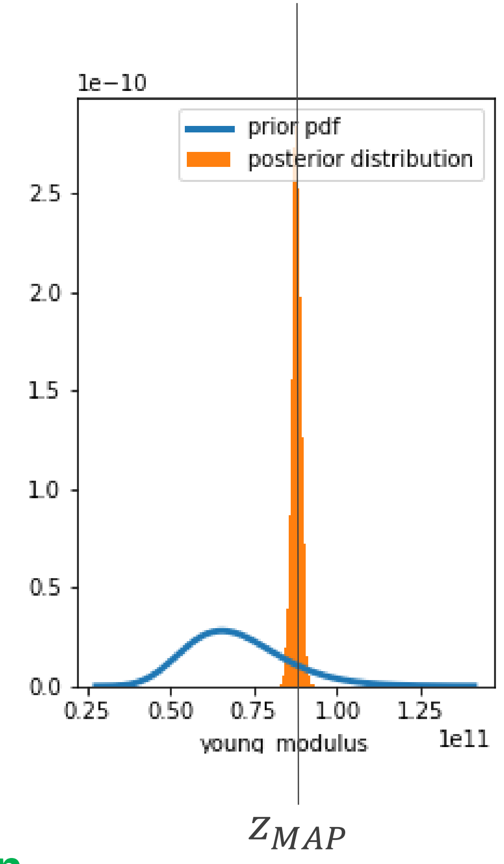
$$p(\, Y_e \mid Z \,) \sim \mathcal{N}(Y(Z), \sigma^2 \mathbb{I})$$

- **Bayes theorem yields**

$$\log p(\, Z \mid Y_e = y_e \,) = -\frac{1}{2} \frac{\|y_e - Y(Z)\|^2}{\sigma^2} + \log p(Z) - \log p(Y_e = y_e)$$

- **We can therefore compute the MAP estimate by solving the optimization problem**

$$z_{MAP} \in \arg\max_{z} - \underbrace{\frac{1}{2} \frac{\|y_e - Y(z)\|^2}{\sigma^2}}_{\text{Least square problem}} + \underbrace{\log p(Z)}_{\text{with a penalty function}}$$



STEP B': Variational inference

- **Objective**
  - Approximate the posterior distribution by a variational distribution $q_\theta(Z)$ with parameters θ

$$q_\theta(Z) \approx p(\, Z \mid Y_e = y_e \,)$$

- **Methodology**
  - **Mean field approximation** of the posterior distribution (dependences are lost)

$$q_\theta(Z) = q_\theta^1(Z_1) \ldots q_\theta^d(Z_d)$$

  - Computation of the optimal distribution by minimization of a **Kullback-Leibler divergence** ("distance") between distributions

$$\min_\theta KL(q_\theta(Z) \,\|\, p(\, Z \mid Y_e = y_e \,)) = \mathbb{E}_{q_\theta}\left(\frac{\log q_\theta(Z)}{\log p(Z \mid Y_e = y_e)}\right)$$

  - Minimizing the Kullback-Leibler divergence is equivalent to maximing the **Evidence Lower BOund**

$$\max_\theta \mathcal{L}(q_\theta) = \mathbb{E}_{q_\theta}(\log p(\, Y_e = y_e \mid Z \,) + \log p(Z) - \log q_\theta(Z))$$

STEP B’: ELBO optimization

- **Maximization of the ELBO**

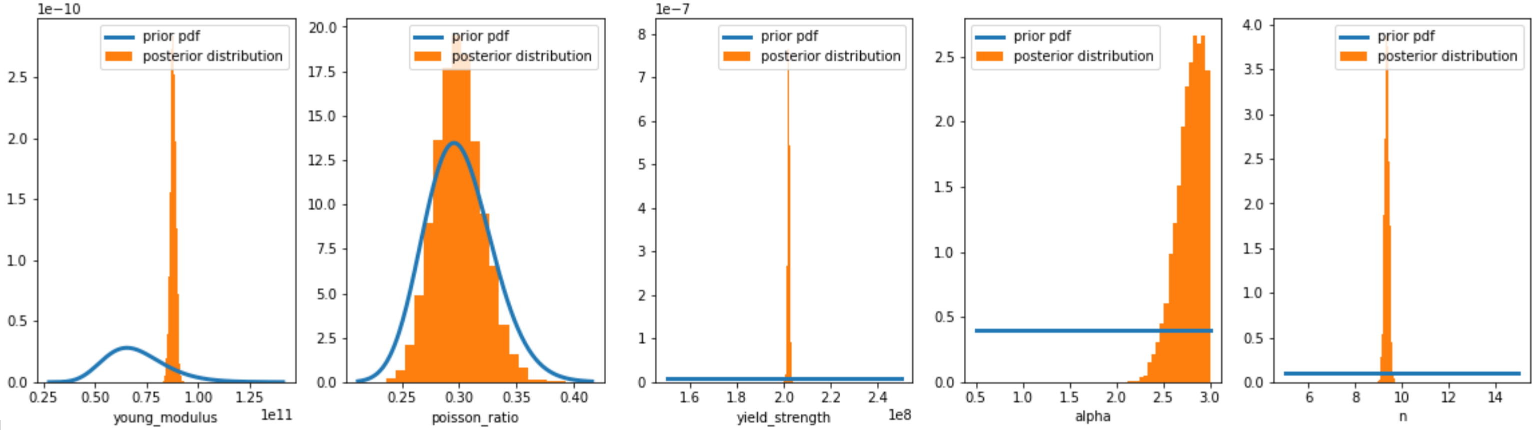
$$\max_{q_\theta} \mathcal{L}(q_\theta) = \mathbb{E}_q(\log p(Y_e = y_e \mid Z) + \log p(Z) - \log q_\theta(Z))$$

- **Monte-Carlo approximation of the ELBO and its gradient**

$$\mathcal{L}(q_\theta) \approx \frac{1}{N}\sum_{i=1}^{N} \log p(Y_e = y_e \mid z_i) + \log p(z_i) - \log q_\theta(z_i)$$

$$\nabla_\theta \mathcal{L}(q_\theta) \approx \frac{1}{N}\sum_{i=1}^{N} (\nabla_\theta \log q_\theta(z_i))(\log p(Y_e = y_e \mid z_i) + \log p(z_i) - \log q(z_i))$$

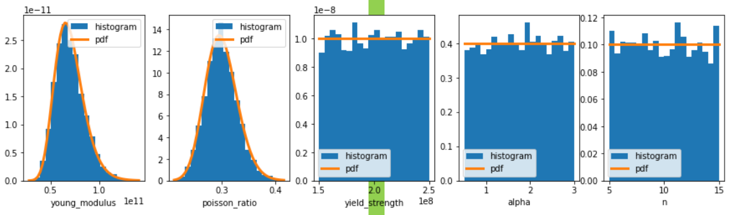
- **Optimization by stochastic gradient ascent – Black-box Variational Inference algorithm**
  - Initialize $\theta, \rho > 0, N$
  - While θ has not converged:
    - Sample $(z_i)_{i=1}^{N}$ according to q_θ
    - Compute the Monte-Carlo approximation $\Delta\theta$ of $\nabla_\theta \mathcal{L}(q_\theta)$
    - $\theta \leftarrow \theta + \rho\Delta\theta$
  - Return q_θ

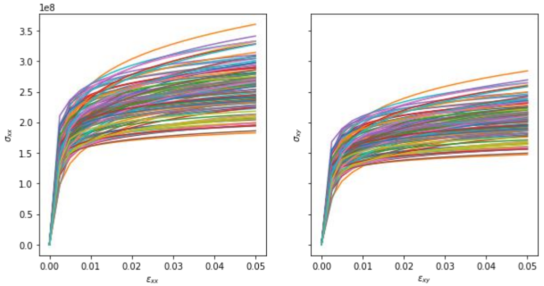


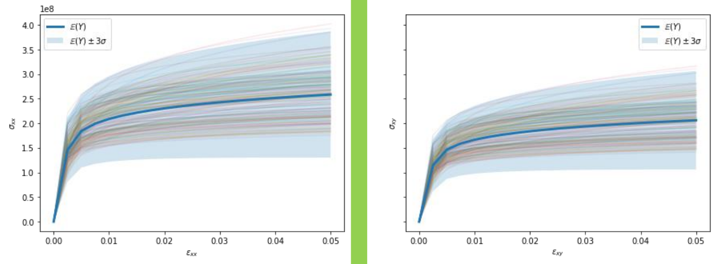
STEP B
Uncertainty quantification
Modeling uncertainty sources with probability distribution

STEP A
Problem specification
Computational model definition
Assessment criteria definition
$Y = M(X)$

STEP C
Uncertainty propagation







STEP C'
Sensitivity analysis
Ranking of the sources of uncertainty

STEP B'
Quantification of sources
Inverse methods, Bayesian inference

Many thanks for your attention