An Industrial Viewpoint on Uncertainty Quantification in Simulation: Stakes, Methods, Tools, Examples

Alberto Pasanisi
Project Manager

EDF R&D. Industrial Risk Management Dept.
Chatou, France

alberto.pasanisi@edf.fr

EDF
LEADING THE ENERGY CHANGE

Summary

- Common framework for uncertainty management
- Examples of applied studies in different domains relevant for EDF :
    - Nuclear Power Generation
    - Hydraulics
    - Mechanics

Common framework for uncertainty management

Which uncertainty sources?

- The modeling process of a phenomenon contains many sources of uncertainty:
  - model uncertainty: the translation of the phenomenon into a set of equations. The understanding of the physicist is always incomplete and simplified,
  - numerical uncertainty: the resolution of this set of equations often requires some additional numerical simplifications,
  - parametric uncertainty: the user feeds in the model with a set of deterministic values ... According to his/her knowledge

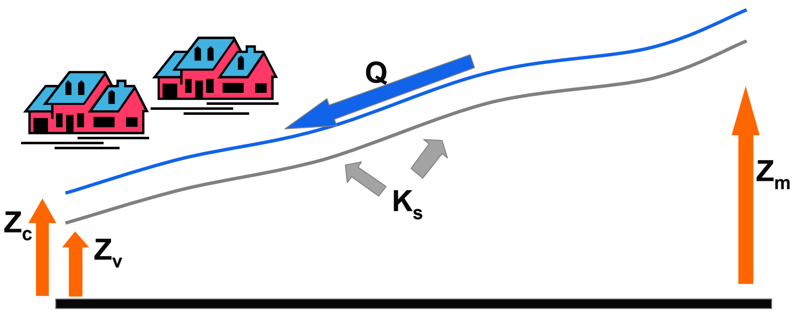
- Different kinds of uncertainties taint engineering studies; we focus here on parametric uncertainties (as it is common in practice)

Which (parametric) uncertainty sources?

- Epistemic uncertainty
  - It is related to the lack of knowledge or precision of any given parameter which is deterministic in itself (or which could be considered as deterministic under some accepted hypotheses). E.g. a characteristic of a material.
- Stochastic (or aleatory) uncertainty
  - It is related to the real variability of a parameter, which cannot be reduced (e.g. the discharge of a river in a flood risk evaluation). The parameter is stochastic in itself.
- Reducible vs non-reducible uncertainties
  - Epistemic uncertainties are (at least theoretically) reducible
  - Instead, stochastic uncertainties are (in general) irreducible (the discharge of a river will never be predicted with certainty)
    - A counter-example: stochastic uncertainty tainting the geometry of a mechanical piece → Can be reduced by improving the manufacturing line … The reducible aspect is quite relative since it depends on whether the cost of the reduction actions is affordable in practice

A (very) simplified example

Flood water level calculation



Uncertainty

$$Z_c = Z_v + \left[\frac{Q}{K_s \cdot \sqrt{(Z_m - Z_v)/L} \cdot B} \right]^{3/5}$$

Strickler's Formula

- Z_c : Flood level (variable of interest)
- Z_m et Z_v : level of the riverbed, upstream and downstream (random)
- Q : river discharge (random)
- K_s : Strickler's roughness coefficient (random)
- B, L : Width and length of the river cross section (deterministic)

General framework

Input Variables	***Model***	***Output variables of interest***
Uncertain : *X* Fixed : *d*	*G(X,d)*	*Z = G(X, d)*

Which output variable of interest?

- Formally, we can link the output variable of interest *Z* to a number of continuous or discrete uncertain inputs *X* through the function G:

$$Z = G(X, d)$$

  - *d* denotes the "fixed" variables of the study, representing, for instance a given scenario. In the following we will simply note: $$Z = G(X)$$

- The dimension of the output variable of interest can be 1 or >1
- Function G can be presented as:
  - an analytical formula or a complex finite element code,
  - with high / low computational costs (measured by its CPU time),

- The uncertain inputs are modeled thanks to a random vector *X*, composed of *n* univariate random variables (X_1, X_2, …, X_n) linked by a dependence structure.

Which goal?

- Four categories of industrial objectives:
  - Industrial practice shows that the goals of any quantitative uncertainty assessment usually fall into the following four categories:
    - Understanding: to understand the influence or rank importance of uncertainties, thereby guiding any additional measurement, modeling or R&D efforts.
    - Accrediting: to give credit to a model or a method of measurement, i.e. to reach an acceptable quality level for its use.
    - Selecting: to compare relative performance and optimize the choice of a maintenance policy, an operation or design of the system.
    - Complying: to demonstrate the system's compliance with an explicit criteria or regulatory threshold (e.g. nuclear or environmental licensing, aeronautical certification, ...)
  - There may be several goals in any given study or along the time: for instance, importance ranking may serve as a first study in a more complex and long study leading to the final design and/or the compliance demonstration

Which criteria?

- Different quantities of interest
  - These different objectives are embodied by different criteria upon the output variable of interest.

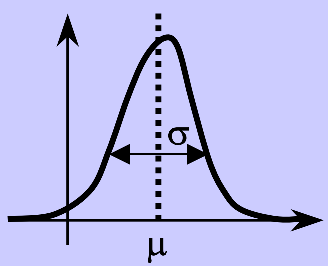
- These criteria can focus on the outputs':
  - range
  - central dispersion
  - "central" value: mean, median
  - probability of exceeding a threshold : usually, the threshold is extreme. For example, in the certification stage of a product.

- Formally, the quantity of interest is a particular feature of the pdf of the variable of interest Z

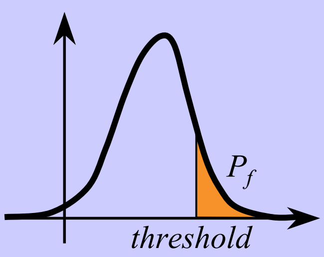
Why are these questions so important?

- The proper identification of:
  - the uncertain input parameters and the nature of their uncertainty sources,
  - the output variable of interest and the goals of a given uncertainty assessment,
- is the key step in the uncertainty study, as it guides the choice of the most relevant mathematical methods to be applied

What is **really** relevant in the uncertainty study?



Mean, median, variance, (moments) of *Z*



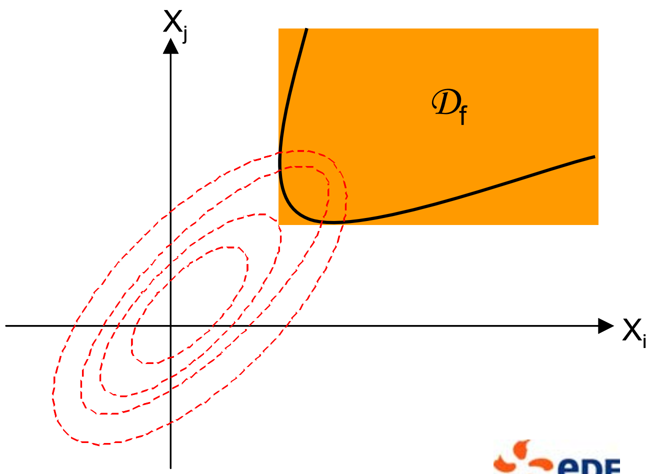
(Extreme) quantiles, probability of exceeding a given threshold

A particular quantity of interest: the “probability of failure”

- G models a system (or a part of it) in operative conditions
  - Variable of interest *Z* → a given state-variable of the system (e.g. a temperature, a deformation, a water level etc.)
- Following an “operator’s” point of view
  - The system is in safe operating condition if *Z* is above (or below) a given “safety” threshold
- System “failure” event: $Z \leq 0$
  - Classical formulation (no loss of generality) in which the threshold is 0 and the system fails when Z is negative
  - Structural Reliability Analysis (SRA) “vision”: Failure if $C\text{-}L \leq 0$ (Capacity – Load)
- Failure domain: $\mathcal{D}_f = \{x \in \mathcal{X} : G(x) \leq 0\}$
- Problem: estimating the mean of the random variable “failure indicator”:

$$I_{\mathcal{D}_f}(x) = \mathbb{1}_{\{G(x)\leq 0\}}$$

$$p_f = \int_{\mathcal{D}_f} f(x)dx = \int_{\mathcal{X}} I_{\mathcal{D}_f}(x)\ f(x)dx = \mathbb{E}\left[I_{\mathcal{D}_f}(X)\right]$$



Need of a generic and shared methodology

- There has been a considerable rise in interest in many industries in the recent decade

- Facing the questioning of their control authorities in an increasing number of different domains or businesses, large industrial companies have felt that domain-specific approaches are no more appropriate.

- In spite of the diversity of terminologies, most of these methods share in fact many common algorithms.

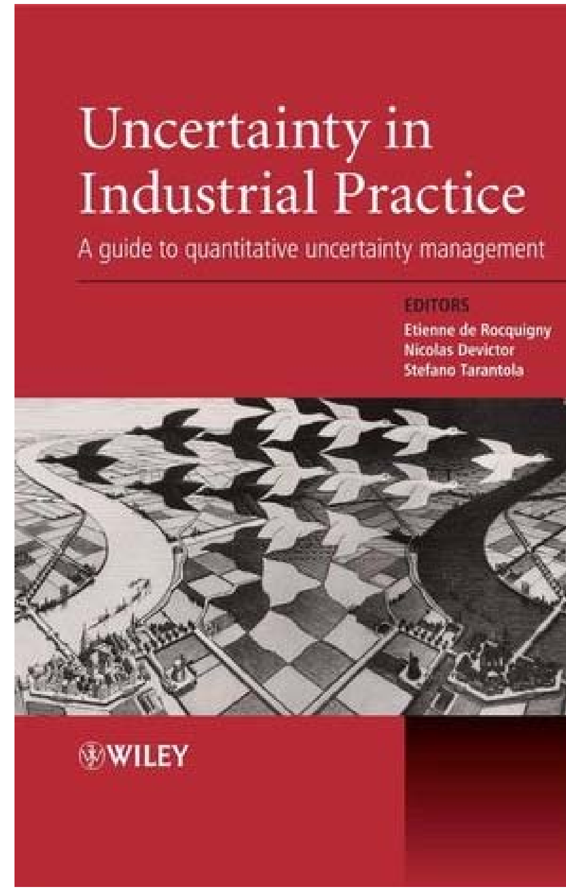
- That is why many industrial companies and public establishments have set up a common methodological framework which is generic to all industrial branches. This methodology has been drafted from industrial practice, which enhances its adoption by industries.

Shared global methodology

The global “uncertainty” framework is shared between EDF, CEA and several French and European partners (EADS, Dassault-Aviation, CEA, JRC, TU Delft …)

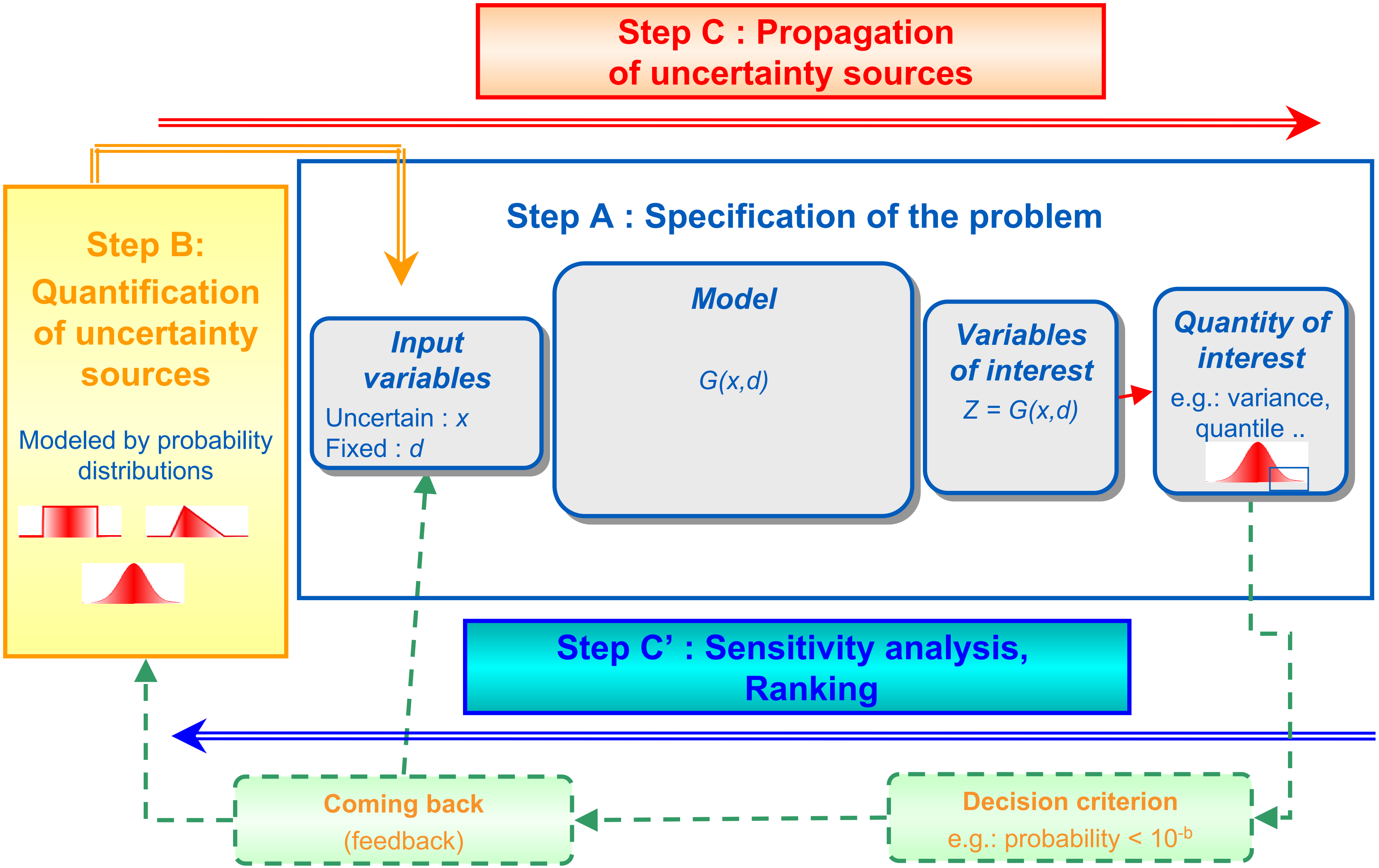
Uncertainty handbook
(ESReDA framework, 2005-2008)

ESReDA
European Safety, Reliability & Data Association



eDF

Uncertainty management - the global methodology



Some comments (Step B). Available information

- Different context depending on the available information
  - Scarce data (or not at all) → Formalizing the expert judgment
    - A popular method: the maximum entropy principle → Between all pdf complying with expert information, choosing the one that maximizes the statistical entropy :

$$H(X) = -\int_{\mathcal{X}} f(x)\log\left(f(x)\right)dx$$

Measure of the "vagueness" of the information on X provided by f(x)

Information	Maximum Entropy pdf
$X \in [a,b]$	Uniform $X \sim \mathcal{U}(a,b)$
$\mathbb{E}(X) = \mu$ $X \in [0,\infty[$	Exponential $X \sim \mathcal{E}(1/\mu)$
$\mathbb{E}(X) = \mu$ $\mathbb{V}(X) = \sigma^2$	Normal $X \sim \mathcal{N}(\mu,\sigma)$

    - Another popular choice: Triangular distribution (range + mode)

  - Feedback data available → Statistical fitting (parametric, non-parametric) in a frequentist or Bayesian framework

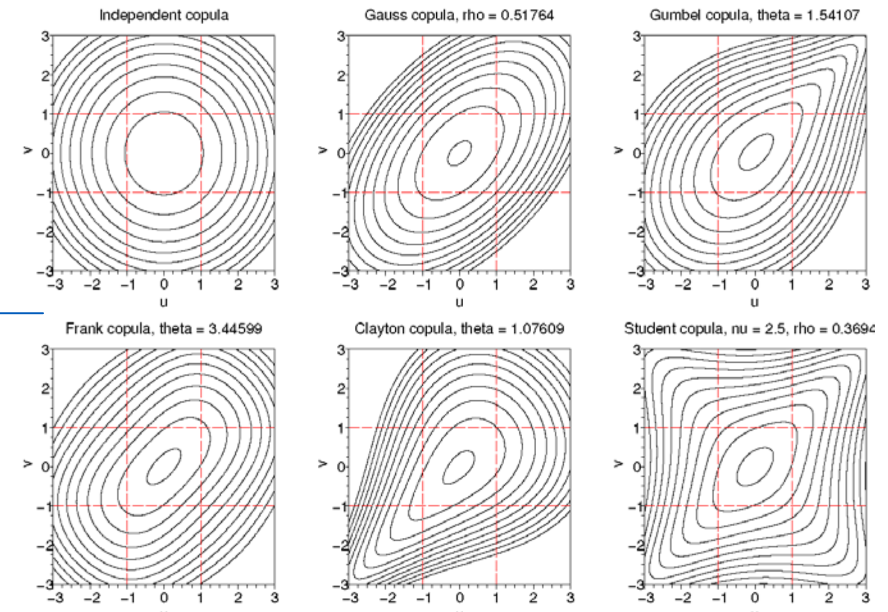
eDF

Some comments (Step B). Dependency

- Taking into account the dependency between inputs is a crucial issue in uncertainty analysis
  - Using copulas structure → CDF of the vector *X* as a function of the marginal CDF of $X_1 \dots X_n$:

$$F(x_1, x_2, \cdots, x_n) = C\left(F(x_1), F(x_2), \cdots, F(x_n)\right)$$

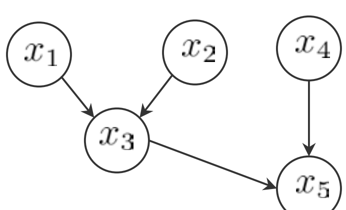
Example: All bivariate densities here have the same marginal pdf's (standard Normal) and the same Spearman rank coeff. (0.5)



  - Using conditional distributions
    - often based on “causality” considerations

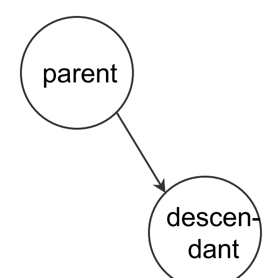
$$f(x_1, x_2) = f(x_1) \cdot f(x_2 | x_1)$$

    - Directed Acyclic Graphs (Bayesian Networks) are helpful for representing the dependency structure



$$f(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} \left(f(x_i | \mathrm{pa}(x_i))\right)$$

Set of the “parents” of x_i

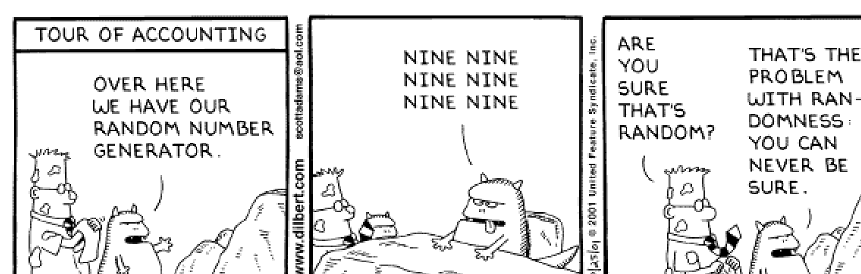


Some comments (Step C and C'). CPU time

- Main issue in the industrial practice: the computational burden!
  - In most problems, the "cost" depends on the number of runs of the deterministic "function" G

- If the code G is CPU time consuming
  - Be careful with Monte-Carlo simulations!
  - Rule of thumb: for estimating a rare probability of 10^{-r}, you need 10^{r+2} runs of G !
  - Appropriate methods (advanced Monte Carlo, meta-modeling)
  - Appropriate software tools for: →
    - Effectively linking the deterministic model G(X) and the probabilistic model F(X)
    - Perform distributing computations (High Performance Computing)
  - Avoid DIY solutions !

www.openturns.org



Examples. Nuclear Power Generation

Nuclear production at EDF

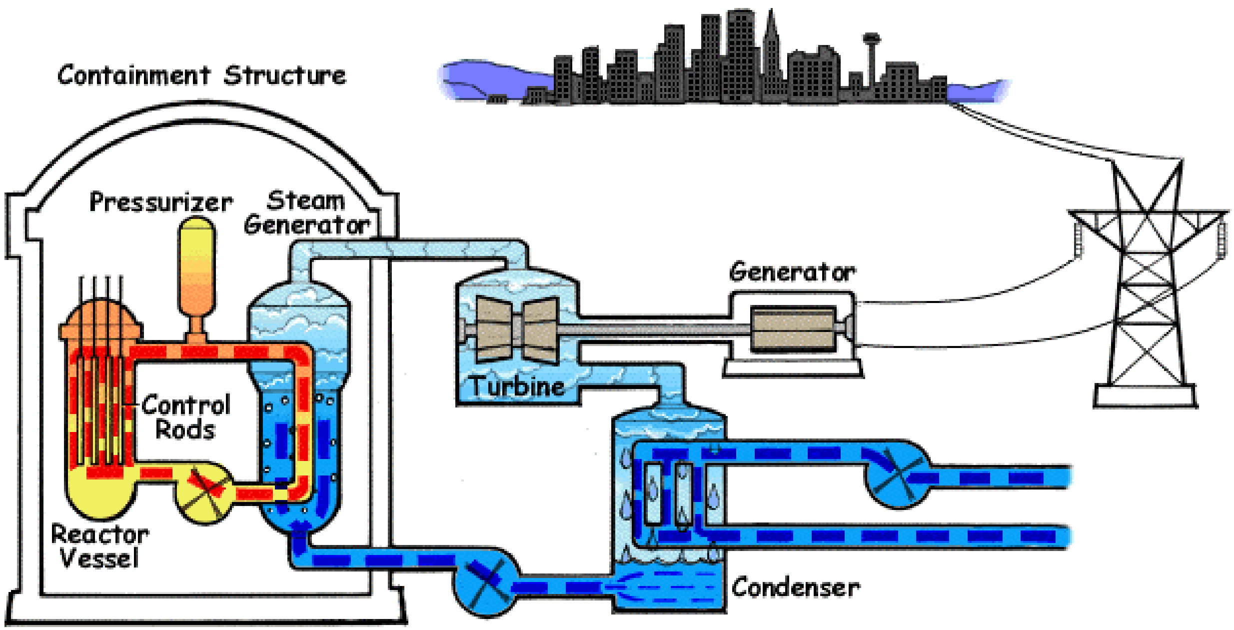
- 58 operating nuclear units in France, located in 19 power stations

- PWR (Pressurized water reactor) technology
  - 3 power levels

- Installed power: 63.1 GW

- Thanks to standard technologies and exploiting conditions, a feedback of more than 1000 operating years

PWR Power unit principles



- **Two separate loops:**
  - **Primary (pressurized water)**
  - **Secondary (steam production)**
- **Three safety barriers (fuel beams, vessel, containment structure)**
- **Highly important stakes**
  - **in terms of safety**
  - **In terms of availability: 1 day off = about 1 M€**

The nuclear reactor pressure vessel (NRPV)

- A key component
- Height: 13 m, Internal diameter: 4 m, thickness: 0,2 m, weight: 270 t
- Contains the fuel bars
- Where the thermal exchange between fuel bars and primary fluid takes place
- It is the second “safety barrier”
- It cannot be replaced !
  - Nuclear Unit Lifetime $\leq$ Vessel Lifetime

- Extremely harsh operating conditions
  - Pressure: 155 bar
  - Temperature: 300 °C
  - Irradiation effects: the steel of the vessel becomes progressively brittle, increasing the risk of failure during an accidental situation

NRPV Safety assessment: a particular UQ problem

- The problem formulation is typical in most nuclear safety problems:
  - **Given some hard (and indeed very rare) accidental conditions**, what is the "failure probability" of the component?
  - It is the case of "structural reliability analysis" (SRA)
  - The physical phenomenon is described by a computer code

The system is safe if Z is lower (or greater) than a fixed value (equal to zero, without loss of generality)

State variable of the system → $Z = G(X)$

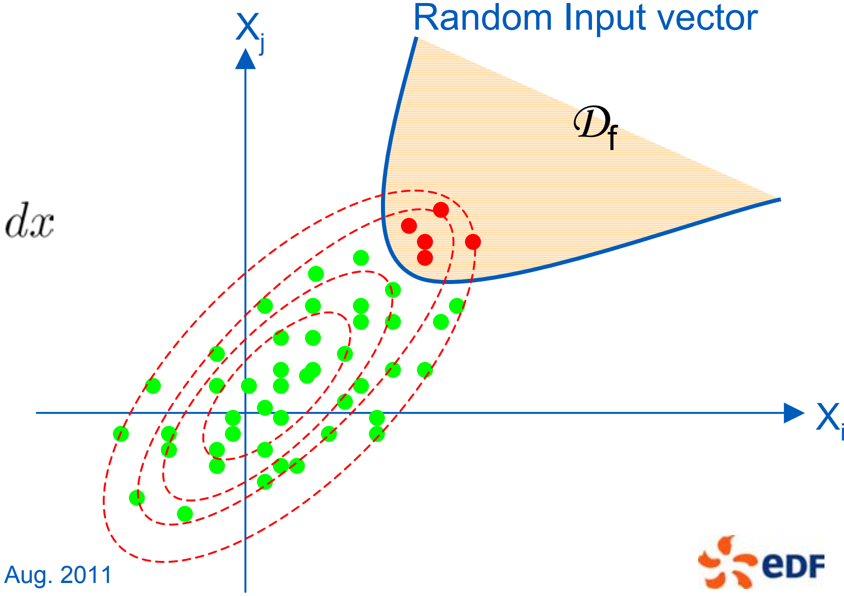
↑ Random Input vector

  - Failure condition: $Z \leq 0$
  - Failure probability

$$P_f = \mathbb{P}(Z \leq 0) = \int_{\mathcal{X}} \mathbb{1}_{\{G(x) \leq 0\}} f(x) dx = \int_{\mathcal{D}_f} f(x) dx$$

$$\mathcal{D}_f = \{x \in \mathcal{X}, G(x) \leq 0\}$$

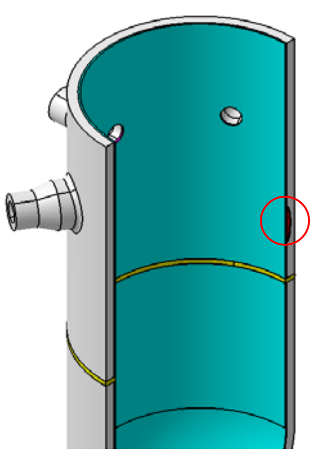
↑ Domain of failure

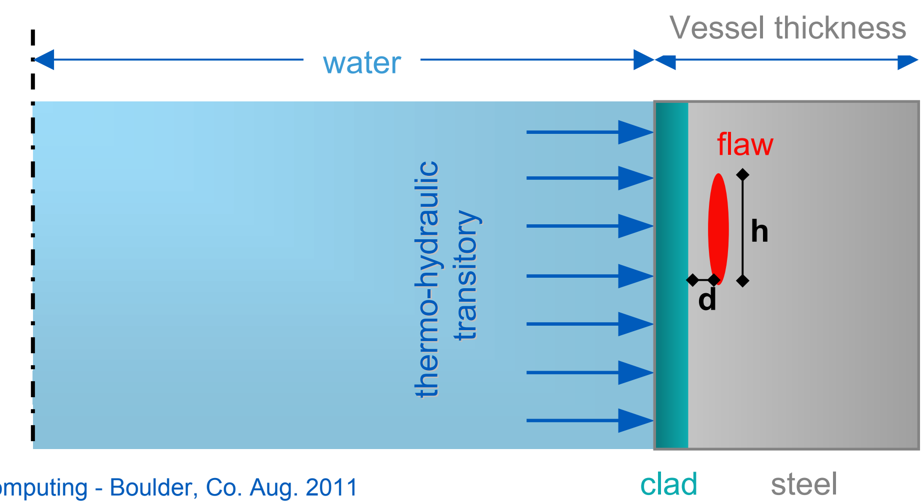


NRPV Safety assessment example [Munoz-Zuniga et al., 2009] (1/3) Step A

- Accidental conditions scenario: cooling water (about 20 °C) is injected into the vessel, to prevent over-warming
  - → Thermal cold shock → Risk of fast fracture around a manufacturing flaw

- Thermo-mechanical fast fracture model:
  - thermo-hydraulic representation of the accidental event (cooling water injection, primary fluid temperature, pressure, heat transfer coefficient)
  - thermo-mechanical model of the vessel cladding thickness, incorporating the vessel material properties depending on the temperature t
  - a fracture mechanics model around a manufacturing flaw
  - Outputs: Stress Intensity $K_{CP}(t)$ in the most stressed point
    Steel toughness, $K_{IC}(t)$ in the most stressed point
  - Goal: Evaluate the probability that for at least one t, the function $G = K_{IC} - K_{CP}$ is negative





NRPV Safety assessment example [Munoz-Zuniga et al., 2009] (2/3) Step B

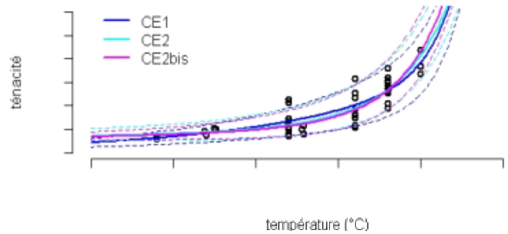
- A huge number of physical variables …
- In this example, three are considered as random. Penalized values are given to the remaining variables

Variable	Distribution	Distribution parameters	Comments
$u_{K_{Ic}}$	Normal	K_{Ic}^{RCC-M} and variation coefficient : $c_{K_{Ic}} = 15\%$	Support truncated at $[-4\sigma; +4\sigma]$
h	Weibull	Scale parameter $\alpha =$ 3.09 mm and shape $\beta = 1.08$ mm	Distribution estimated by fitting exercise over inspection data
d	Uniform	$[0.1; 100]$ (mm)	The flaw is supposed to be in the inner half-thickness

1) Toughness low limit, playing in the steel toughness law $K_{IC}(t)$
Normal dispersion around a reference value K_{IC}^{RCC}

2) Dimension of the flaw h,

3) Distance between the flaw and the interface steel-clad d,



- A more complex example with 7 randomized inputs is given in [Munoz-Zuniga et al., 2010]

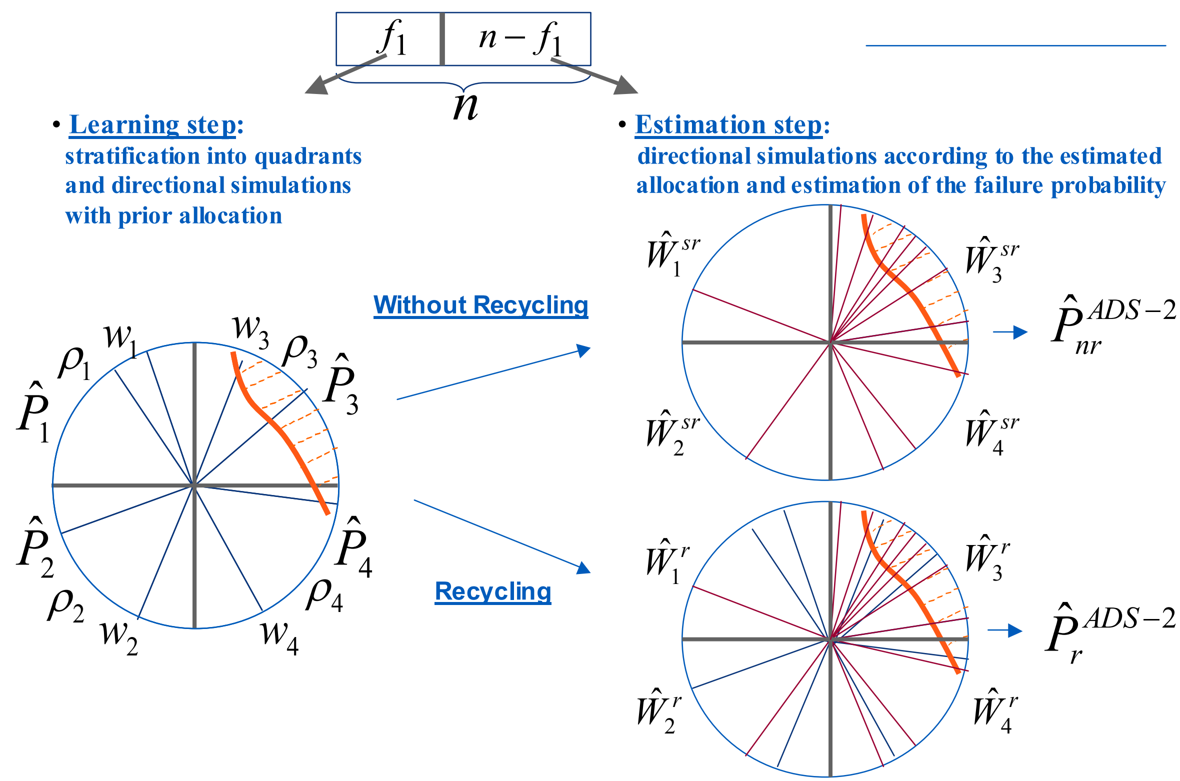
NRPV Safety assessment example [Munoz-Zuniga et al., 2009] (3/3) Step C

- A numerical challenge:
  - High CPU time consuming model
  - Standard Monte Carlo Methods are inappropriate to give an accurate estimate of P_f
  - An innovative Monte Carlo sampling strategy has been developed: “ADS-2” (Adaptive Directional Stratification)

- A numerical challenge:
  - Standard transformation
  - Directional sampling
  - Adaptive strategy to sample more “useful” directions

	Method	p	n	$f_1(n)$	$\hat{P}_f$	95% CI length	Nb. of calls to G_{Min}
1	DS	3	50	/	4.3×10^{-6}	1.3×10^{-5}	208
2	2-ADS	3	50	10	6.3×10^{-6}	7.0×10^{-6}	820
3	DS	3	200	/	3.7×10^{-6}	1.6×10^{-5}	822
4	2-ADS	3	200	40	8.0×10^{-6}	5.0×10^{-6}	3290
5	DS	3	1000	/	7.66×10^{-6}	7.0×10^{-6}	4028

Example of results. **NB P_f is here conditional to the occurrence of very rare accidental conditions**



Examples. Hydraulics

Hydraulic simulation: a key issue

- Hydraulic simulation is a key issue for EDF
- Because EDF is a major hydro-power operator
  - mean annual production: 40 TWh
  - 220 dams, 447 hydro-power stations
- Because (sea or river) water plays a key role in nuclear production

Example of UQ in hydraulic simulation: effects of the embankment's failure hydrograph on flooded areas assessment [Arnaud et al., 2010] (1/5)

Context: French regulations for large dams

- Large dams are considered as potential sources of major risks (Law 22/07/1987)
- Emergency Response Plans (PPI) must be prepared by the local authority ("Préfet") after consultation
- Risk assessment study :
  - Risk assessment **in case of dam failure**: Evaluation of the **Maximum water level** (Z_{max}) and **wave front arrival time** (T_{front})
  - Seismic analysis
  - Evaluation of the possibility and effect of landslide in the reservoir
  - Hydrology study

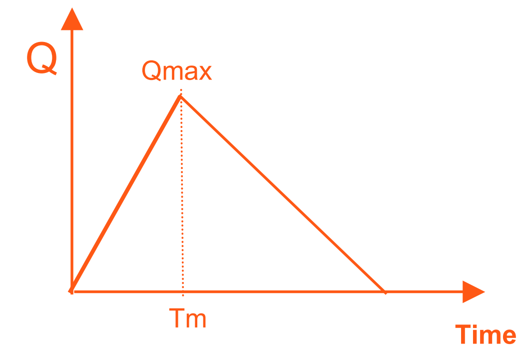
Hypotheses for the dam failure:

- Concrete dams : the dam collapses instantaneously
- Earth dams : the dam failure is assumed to be progressive by the formation of a breach due to internal erosion or an overflow → Embankment failure hydrograph

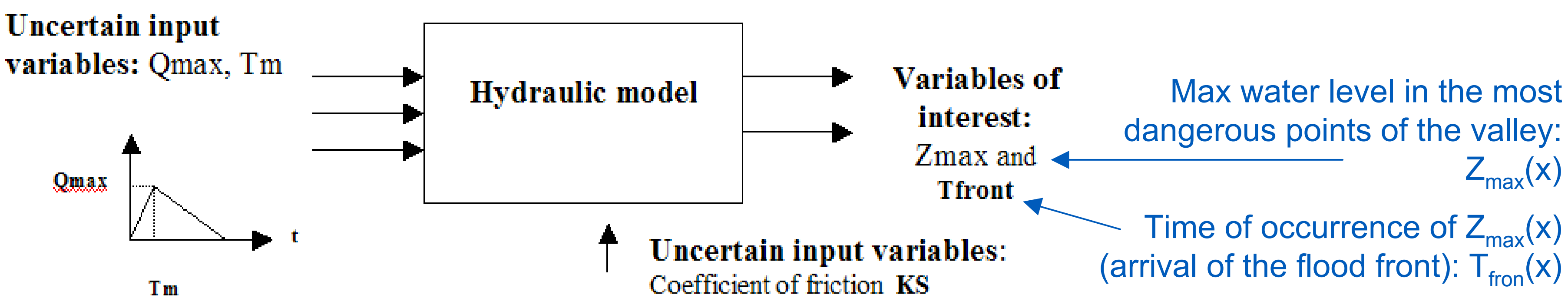
Effects of the embankment's failure hydrograph on flooded areas assessment [Arnaud et al., 2010] (2/5)

- The complex physics at play during the progressive erosion is not well known
  - the emptying hydrograph H is not well known:
    - The maximum discharge Qmax
    - The time of occurrence of the maximum discharge Tmax
  - We assume that the reservoir volume (V) is known
  - We assume a triangular hydrograph

$$H : t \mapsto Q = H(t)$$
$$\int H(t)dt = V$$



- **Step A**



Effects of the embankment's failure hydrograph on flooded areas assessment [Arnaud et al., 2010] (3/5)

- Known variables:
  - Features of the dam
    - Dam height 123 m, Reservoir volume: V=1200 Mm3
  - Valley features
    - Length : 200 km, no tributaries, no dams downstream
    - Very irregular geometry with huge width variation → Hydraulic jumps

Step B → Uncertainty assessment

- Q_{max} and T_{max} (Hydrograph form)
  - too small amount and imprecise data: the pdf could not be assessed by a statistical procedure
  - According to the expert advice the following pdf's for Q_{max} and T_m have been proposed:

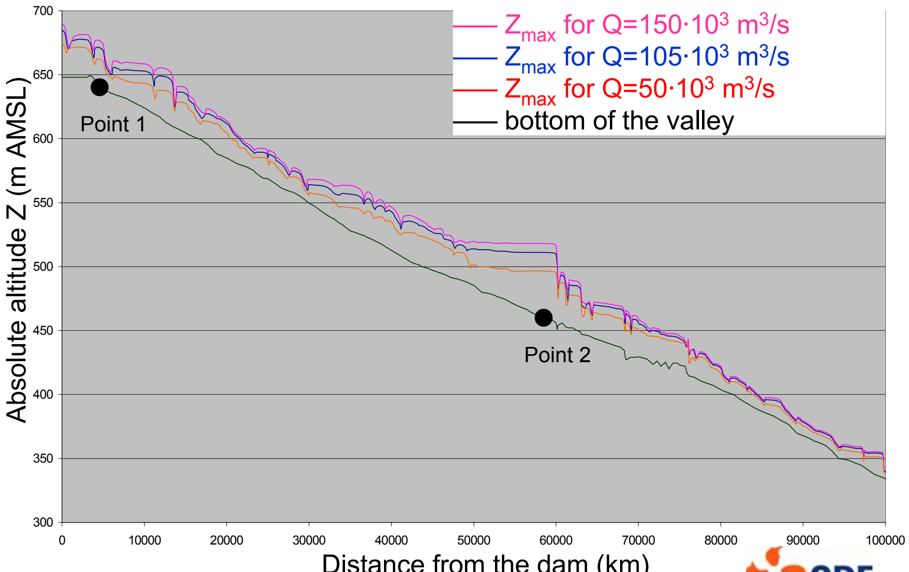
- Friction coefficient Ks
  - Not "measurable" variable
  - Expert advice, based on valley morphology knowledge

Prob. distr. funct	Q_{max} (m3/s)	T_m (s)
1) Normal : Mean Standard dev.	 100 000 25 000	 5 000 2 000
2) Uniform : Lower bound Upper bound	 50 000 150 000	 1 000 7 200

Prob. distr. funct	Ks
1) Truncated Normal: Mean Standard dev. Bounds	 30 5 [17.5, 47.5]
2) Uniform: Lower bound Upper bound	 25 35

Effects of the embankment's failure hydrograph on flooded areas assessment [Arnaud et al., 2010] (4/5)

- **Step C → Uncertainty propagation**
- Hydraulics software: "Mascaret" Code (EDF R&D-CETMEF)
  - 1D shallow water modeling based on the De St Venant equations
  - Finite volume scheme with CFL limitation on the time step
- Hydraulic modeling
  - Un-stationary flow conditions, Space discretization: 100 m
  - The time step (1-2 s) is controlled by the CFL condition. Duration of the simulation : 13 000 time steps
- First set of 3 runs of the model to look for the more dangerous points
  - 3 values of Q_{max} : 50 000 m³/s, 105 000 m³/s and 150 000 m³/s
  - Mean value of Ks
  - Two points (Point 1 and Point 2) are particularly dangerous with respect to the flooding risk. They are both located downstream from a section narrowing → hydraulic jumps
  - We will mainly focus on these two points



Effects of the embankment's failure hydrograph on flooded areas assessment [Arnaud et al., 2010] (5/5)

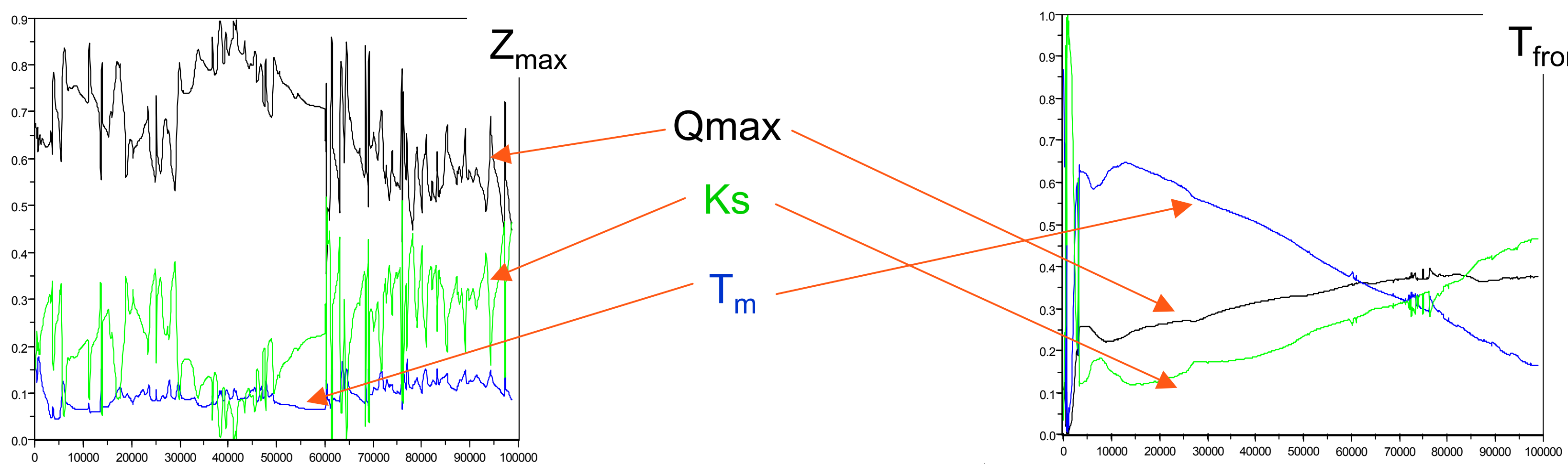
- Propagation method: Surface response + Monte Carlo
- Some results
  - Extreme Quantiles of Z_{max} in points 1 and 2 (flood risk assessment)

Z_{max} (m ASML)	Point 1		Point 2	
	Pdf 1	Pdf 2	Pdf 1	Pdf 2
Quantile 99.9%	676.66	676.64	517.14	515.04
Quantile 99%	675.57	675.52	516.49	515.57
Quantile 95%	673.67	674.25	513.71	514.14

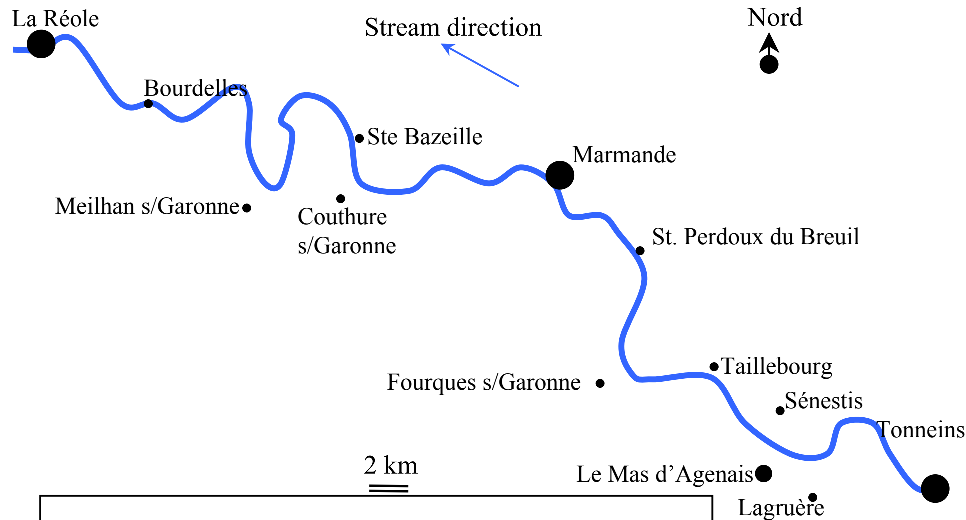
  - Sensitivity analysis → evaluation of the Spearman ranks' correlation coefficients for all values of the abscissa x

$\rho^S(Q_{max}, Z_{max}), \rho^S(T_m, Z_{max}), \rho^S(Ks, Z_{max})$

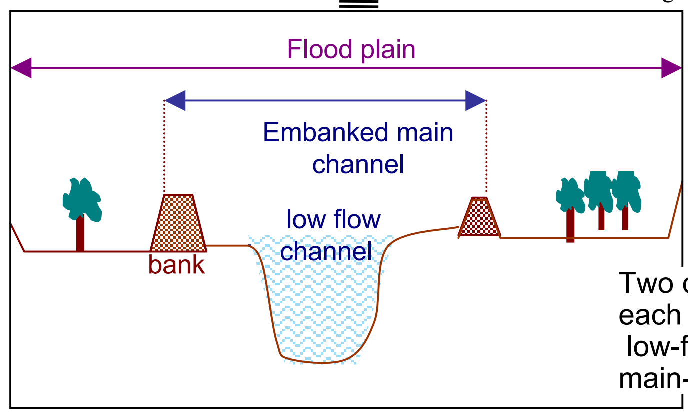
$\rho^S(Q_{max}, T_{fron}), \rho^S(T_m, T_{fron}), \rho^S(Ks, T_{fron})$



A hydraulic benchmark: the Garonne case-study



- Hydraulic modeling of a 50 km long section of the Garonne river → "Mascaret" Code
- Case study shared between the partners of the OPUS project
- Two examples:
  - Inverse modeling to assess the pdf of Strickler's roughness coefficient Ks
    - K_s is never directly observed
    - One should estimate the pdf of K_s, given a set of observed coupled data (discharge,water level)
  - Evaluating an extreme quantile of the flood water level at a given abscissa
  - Or evaluating the probability for the flood water level in a given abscissa to be greater than a threshold value



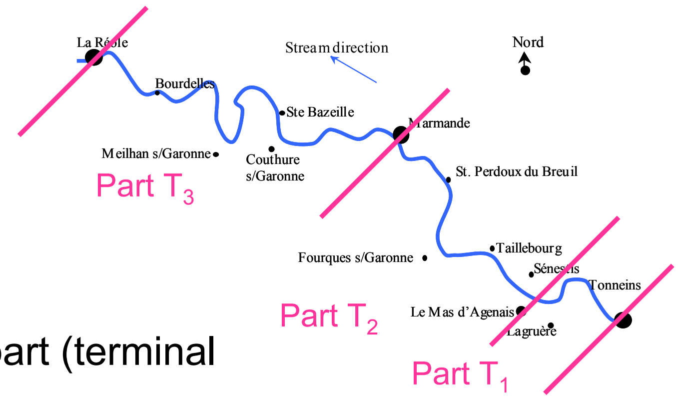
Two different Ks for each section:
low-flow Ks and
main-channel Ks.

AGENCE NATIONALE DE LA RECHERCHE ANR OPUS

"OPen source platform for Uncertainty treatment in Simulation"
10 partners, Tot. budget: 2.2 M€, Leader: EDF

EDF

The Garonne case-study: Inverse modeling of Ks [Couplet, Le Brusquet et al., 2010] (1/2)



- Physical hypothesis
  - 3 parts each one with given values of the 2 Ks

- Statistical problem: assessing the pdf of K_s
  - In this example, we will assess the pdf of the Ks of the T_3 part (terminal part between Marmande and La Réole)
  - Data: couples (discharges Q_i, water levels Z_i) at Mas d'Agenais and Marmande

$$\mathbf{Z}_i = G(Q_i, \mathbf{Ks}) + \mathbf{U}_i, \quad i = 1..n, \quad \mathbf{Z}_i, \mathbf{U}_i \in \mathbb{R}^2$$

  - Hypotheses:

$$\begin{pmatrix} Ks_1 \\ Ks_2 \end{pmatrix} \sim \mathcal{N}(\mu, \Sigma) \qquad U_i \sim \mathcal{N}(0, R), R = \sigma_\epsilon \cdot I_2$$

→ The vector Ks and observation errors are normal

The standard measurement error is σ_ε

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

→ Mean values of Ks

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

→ Covariance matrix of Ks

$$\beta = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$$

Tricky likelihood expression

$$\mathcal{L}(\mathbf{z}; \beta) = \prod_{i=1}^{n} L_i(z_i; \beta)$$

$$L_i(z_i; \beta) = \iint_{\mathbf{Ks}} f_{Ks,\beta}(\mathbf{Ks}) \cdot f_\epsilon\left(z_i - G(Q_i, \mathbf{Ks})\right) d\mathbf{Ks}$$

Density of **Ks** — Density of $\mathbf{z}_i$, given Q_i and **Ks**

EDF

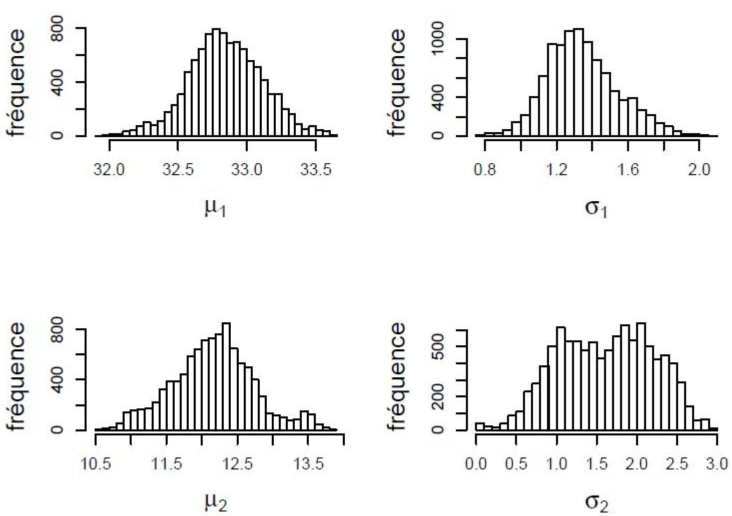
The Garonne case-study: Inverse modeling of Ks [Couplet, Le Brusquet et al., 2010] (2/2)

- Some results
- Two solutions
  - Likelihood maximization (variants of the EM algorithm: ECME, SAEM)

$$\hat{\beta}_{\text{ECME}} = (32.84, 12.03, 0.84, 1.33, 0.35)$$

$$\hat{\beta}_{\text{SAEM}} = (32.78, 12.12, 0.85, 1.48, 0.18)$$

  - Bayesian solution: MCMC sampling from the posterior pdf of β: $\pi(\beta|\mathbf{z}) \propto \pi_0(\beta) \cdot \mathcal{L}(\mathbf{z};\beta)$

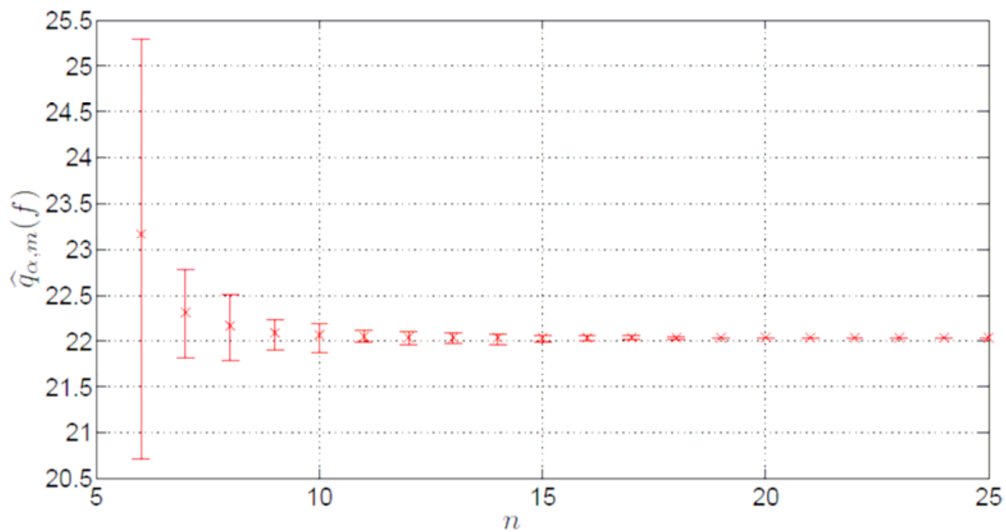


NB: Uniform prior used

The Garonne case-study: Flood risk assessment [Arnaud, Vazquez, Bect et al., 2010]

- Goal: Evaluating the quantile of probability α=0.99 of the water level in a given section
  - Original meta-modeling technique developed within the OPUS project [Vazquez et al, 2010]
  - Empirical estimation of the quantile: $\hat{q}_{\alpha,m} = z^{\lceil m\alpha \rceil}$
  - Building an approximation $\tilde{G}_n(\cdot)$ of $G(\cdot)$ based on the n<<m evaluations: $\{G(x_1), G(x_2), ..., G(x_n)\}$
  - The n points $\{x_1, x_2, ..., x_n\}$ are chosen sequentially in order to minimize a statistical "cost" (e.g. a quadratic loss) between $\hat{q}_{\alpha,m}$ and the empirical estimator built according to the surrogate model

With a dozen runs of the model, it is possible to build a "specialized" kriging meta-model for the quantile estimation (here m=2000)



Examples
Mechanics

A longstanding experience at EDF R&D

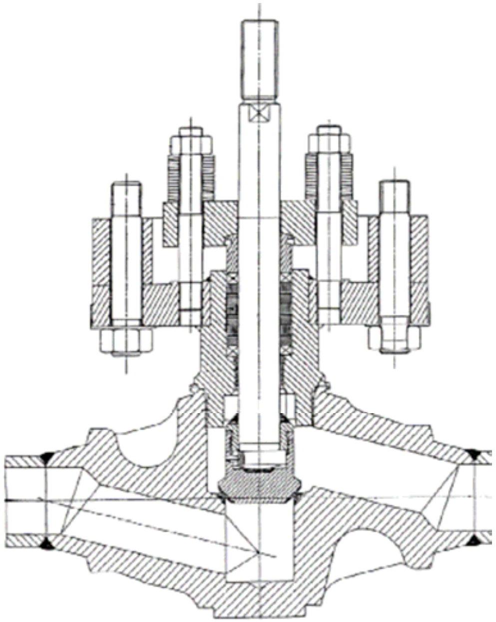
- Several studies in the field of probabilistic mechanics:
  - Reliability analysis
  - Sensitivity analysis
  - Inverse problems → Bayesian updating of the behavior law of the material (e.g. concrete in civil works studies)

- Several research works on polynomial chaos expansion
  - A useful tool to perform high CPU time-consuming calculations above

- Numerous applications
  - Cooling towers, containment structures, thermal fatigue problems, lift-off assessment of fuel rod ...
  - We will focus on an application concerning reliability and sensitivity analysis of globe valves

Globe valve reliability and sensitivity analysis [Berveiller et al., 2010] (1/5)

- Industrial globe valves are used for isolating a piping part inside a circuitry
- Harsh operating conditions: water temperature, pressure, corrosion problems ...
- Reliability assessment: the tightness of the valve has to be assured even with a maximum pressure of the water

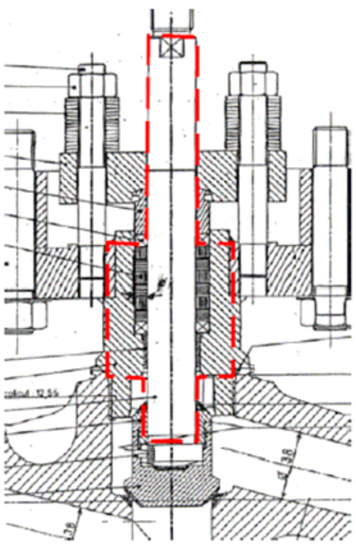


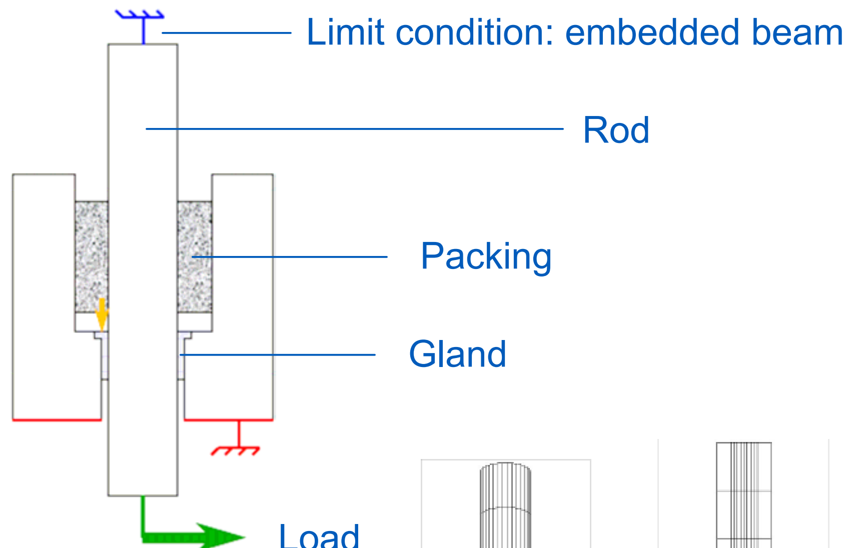
- Several uncertain variables
  - Material properties
  - Functional clearances
  - Load

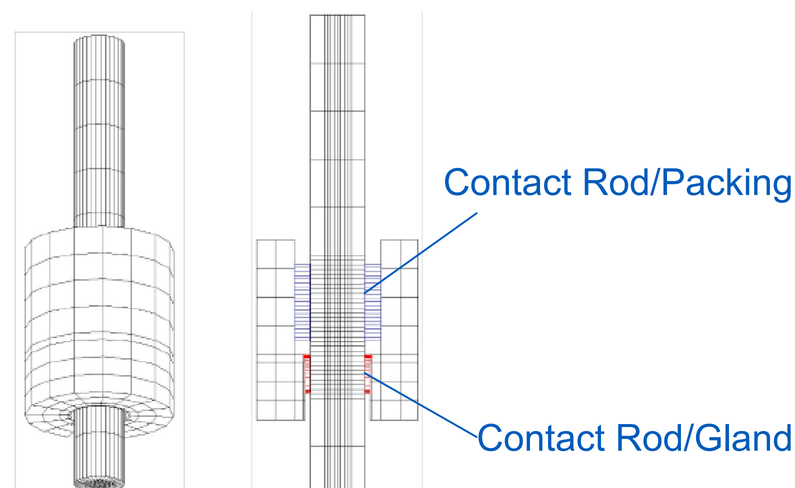
To ensure the reliability of the mechanism, the contact pressures and the max displacement of the rod must be lower than given values

Globe valve reliability and sensitivity analysis [Berveiller et al., 2010] (2/5)

- The modeling problem is very complex. We will work here on a simplified mechanical modeling







Case-study of the OPUS project

Globe valve reliability and sensitivity analysis [Berveiller et al., 2010] (3/5)

Step A

- Variables of interest:
  - Contact pressures
  - Max displacement of the rod
- 6 Uncertain input variables:
  - Packing Young's modulus
  - Gland Young's modulus
  - Beam Young's modulus
  - Steel (Rod) Young's modulus
  - Load
  - Clearance
- Deterministic model G(·):
  - FEM Numerical model of the simplified scheme using *Code_Aster* software (www.code-aster.org)

- Goal of the study:
  - assessing the sensitivity of the variable of interest with respect to the uncertain inputs
- Quantities of interest: Sensitivity indices
  - Reminder: Sobol' variance decomposition*

$$\mathbb{V}[Z] = \sum_i V_i[Z] + \sum_{i<j} \mathbb{V}_{ij}[Z] + \ldots$$

$$V_i[Z] = \mathbb{V}\left[\mathbb{E}[Z|X_i]\right]$$
$$V_{ij}[Z] = \mathbb{V}\left[\mathbb{E}[Z|X_i, X_j]\right] - V_j[Z] - V_j[Z]$$

*X_i's independent

  - Sobol' indices:

$$S_i = \frac{V_i}{\mathbb{V}[Z]} \qquad S_{ij} = \frac{V_{ij}}{\mathbb{V}[Z]} \qquad S_i^T = S_i + \sum_{i \neq j} S_{ij} + \ldots$$

First order — Second order — "Total" index

  - They measure the "part" of the global variance explained by a single input (or a set of inputs)
  - Monte Carlo calculation is CPU expensive, as many model runs are needed → Meta-modeling approach

Globe valve reliability and sensitivity analysis [Berveiller et al., 2010] (4/5)

Step B

- Uncertainty modeling of input variables:

Variable	Prob. density	Mean	Coefficient of Variation
Packing Young's modulus (MPa)	LogNormal	100 000	20%
Gland Young's modulus (MPa)	LogNormal	207 000	10%
Beam Young's modulus (MPa)	LogNormal	6 000	10%
Steel (Rod) Young's modulus (MPa)	LogNormal	200 000	10%
Load (N)	Normal	10 000	10%
Clearance (mm)	$\text{Beta}_{[0,0.1]}$	0.05	50%

Steps C,C'

- Non intrusive polynomial chaos approximation
  - Isoprobabilistic transformation of the input vector: $\xi_i = T(X_i) \sim \mathcal{N}(0,1) \quad i = 1, ..., m$
  - Polynomial chaos (PC) approximation: $Z \approx \widetilde{Z} = \sum_{i=0}^{Q-1} \alpha_j \cdot \psi_j(\boldsymbol{\xi})$ ← PC approx. of order m and degree q

Number of terms of the sum: $Q = \frac{(m+q)!}{m!\, q!}$

coefficients

$\{\psi_j(\boldsymbol{\xi}), j = 1, .., Q\}$ Set of the m-dimensional Hermite polynomials of degree ≤ q

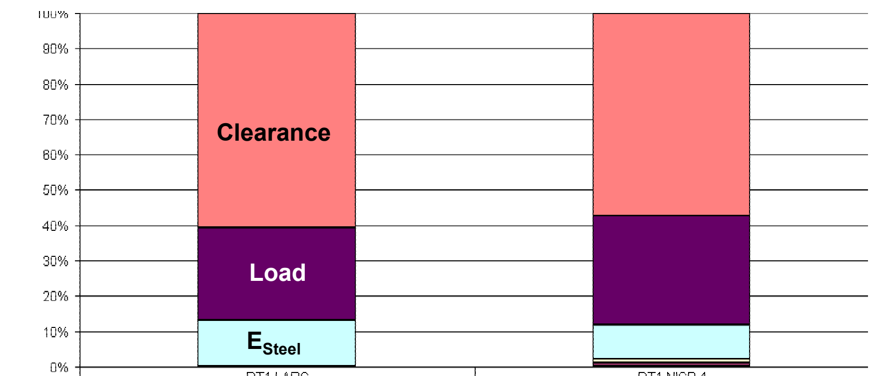
Globe valve reliability and sensitivity analysis [Berveiller et al., 2010] (5/5)

- Benefits of PC approximation
  - Once coefficients are evaluated, PC expansion allows performing quick Monte Carlo simulations, by running the meta-model instead of the expensive numerical code G(·)
  - Moreover, due to the orthogonality of the polynomials, the evaluation of Sobol' indices is straightforward
  - [Sudret, 2008]: $\mathbb{V}[Z] = \sum_{j=1}^{Q-1} \alpha_j^2 \cdot \mathbb{E}\left[\psi_j^2(\boldsymbol{\xi})\right]$ $\qquad V_i[Z] = \mathbb{V}\left[\mathbb{E}[Z|X_i]\right] = \sum_{k \in \mathcal{A}_i} \alpha_k^2 \cdot \mathbb{E}\left[\psi_k^2(\boldsymbol{\xi})\right]$

    Set of polynomials containing only ξ_i
  - The calculation burden (i.e. running several times the code G) is focused on the estimation of the coefficients
    - Several techniques: projection, regression, simulation, sparse PC expansion (LARS) [Blatman & Sudret, 2010]

Example of results

- Sobol' indices for rod displacement
  - PC approximation built by two different methods & tools: LARS, NISP (CEA)
  - Most influent variables : clearance, load, Steel Young's modulus



	DT1 LARS	DT1 NISP 4
JEU	58,92%	54,05%
F_1	25,52%	29,01%
E_acier	12,81%	9,11%
E_poutre	0,11%	0,97%
E_bague	0,04%	0,98%
E_gar	0,02%	0,24%

eDF

Some references

- Munoz-Zuniga, M. Garnier, J. Remy, de Rocquigny E. (2009). *Adaptive Directional Stratification, An adaptive directional sampling method on a stratified space*. ICOSSAR 09, Sept. 2009, Osaka.
- Munoz-Zuniga, M. Garnier, J. Remy, E. de Rocquigny E (2010). *Adaptive Directional Stratification for controlled estimation of the probability of a rare event*. Submitted.
- Arnaud, A. Goutal N., de Rocquigny, E. (2010). *Influence des incertitudes sur les hydrogrammes de vidange de retenue en cas de rupture progressive d'un barrage en enrochements sur les zones inondées en aval*. SimHydro 2010. June 2010, Sophia Antipolis.
- Couplet, M. Le Brusquet, L. Pasanisi, A. (2010). *Caractérisation des coefficients de Strickler d'un fleuve par inversion probabiliste. 42èmes Journées de Statistique*. Mai 2010, Marseille.
- Arnaud, A. Bect, J. Couplet, C. Pasanisi, A. Vazquez E. (2010) *Evaluation d'un risque d'inondation fluviale par planification séquentielle d'expériences*. 42èmes Journées de Statistique . Mai 2010, Marseille.
- Bect, J. Ginsbourger, D. Ling, L. Picheny, V. Vazquez, E. (2010) *Sequential design of computer experiments for the estimation of a probability of failure* . Statistics and Computing (in press)
- de Rocquigny, E. Devictor, N. Tarantola S. (eds.) (2008). *Uncertainty in industrial practice. A guide to quantitative uncertainty management*. Chichester: J. Wiley & Sons.
- Berveiller, M. Blatman, G. Martinez J.M. (2010). *Analyse de sensibilité d'un robinet à soupape à l'aide de développements sur chaos polynomial*. 42èmes Journées de Statistique . Mai 2010, Marseille.
- Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. Reliability Engineering & System Safety 93 964–979.
- Blatman, G. Sudret, B. (2010). *Efficient computation of global sensitivity indices using sparse polynomial chaos expansions.* Reliability Engineering & System Safety, 95(11): 1216-1229

Thank you for your attention