Methodology of Uncertainty Management

A. Dutfoy, AL. Popelin
EDF R&D

'HPC and Uncertainty Treatment
Examples with OpenTURNS and Uranie'
EDF - Airbus Group - PhiMeca – IMACS - CEA

Prace Advanced Training Center
20-22 may 2014

MAISON DE LA SIMULATION

CHANGER L'ÉNERGIE ENSEMBLE

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?

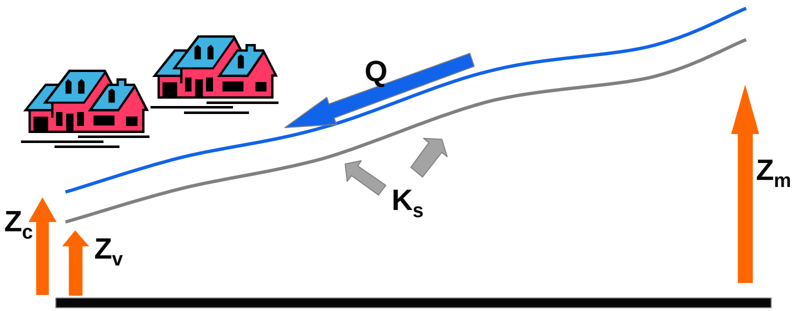
- Epistemic uncertainty
  - It is related to the lack of knowledge or precision of any given parameter which is deterministic in itself (or which could be considered as deterministic under some accepted hypotheses). E.g. a characteristic of a material.

- 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 $(X_1, X_2, \ldots, X_n)$ 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

The major steps in an uncertainty study

- **Step A : Problem's specification**
  - **Input variable(s) - Model - Variable(s) of interest – Quantities of interest**

- Step B : Quantification of uncertainty sources
- Step C : Propagation of uncertainty sources from inputs to outputs
- Step C' : Sensitivity analysis, Ranking uncertainty sources

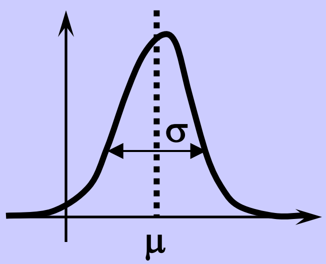
More generic steps (statistics & probability)

Step A : specific to the problem to be solved

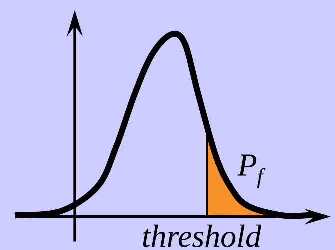
Why are these questions so important?

- 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*

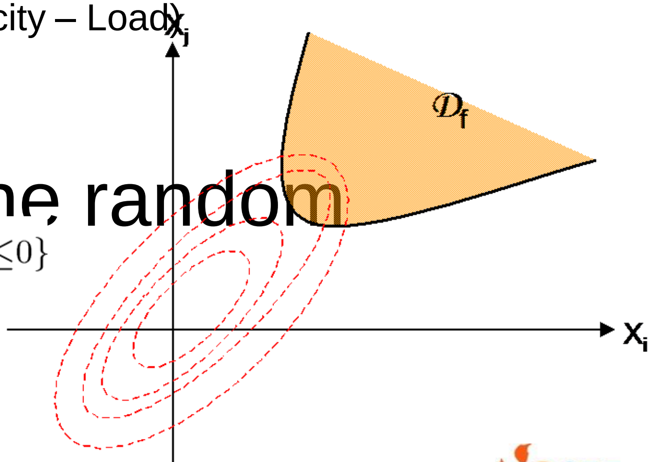


(Extreme) quantiles, probability of exceeding a given threshold

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 $Z \leq 0$
- System "failure" event:
  - 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-$L \leq 0$ (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 of 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.

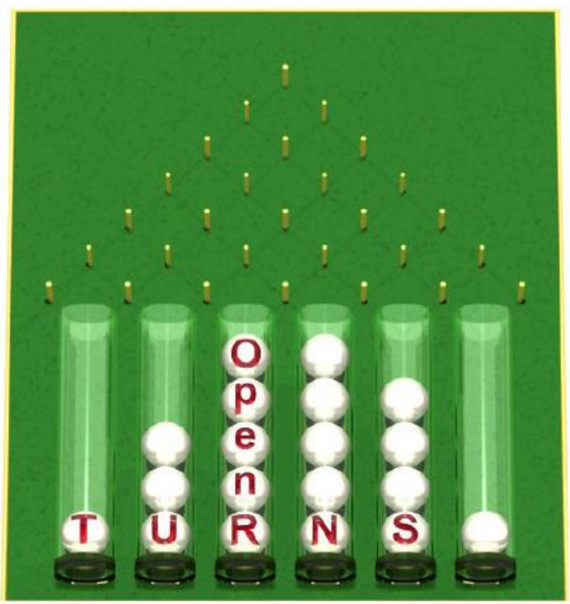
Open TURNS Reference Guide

Reference Guide

OpenTURNS 1.3

Documentation built from package openturns-doc-13.09

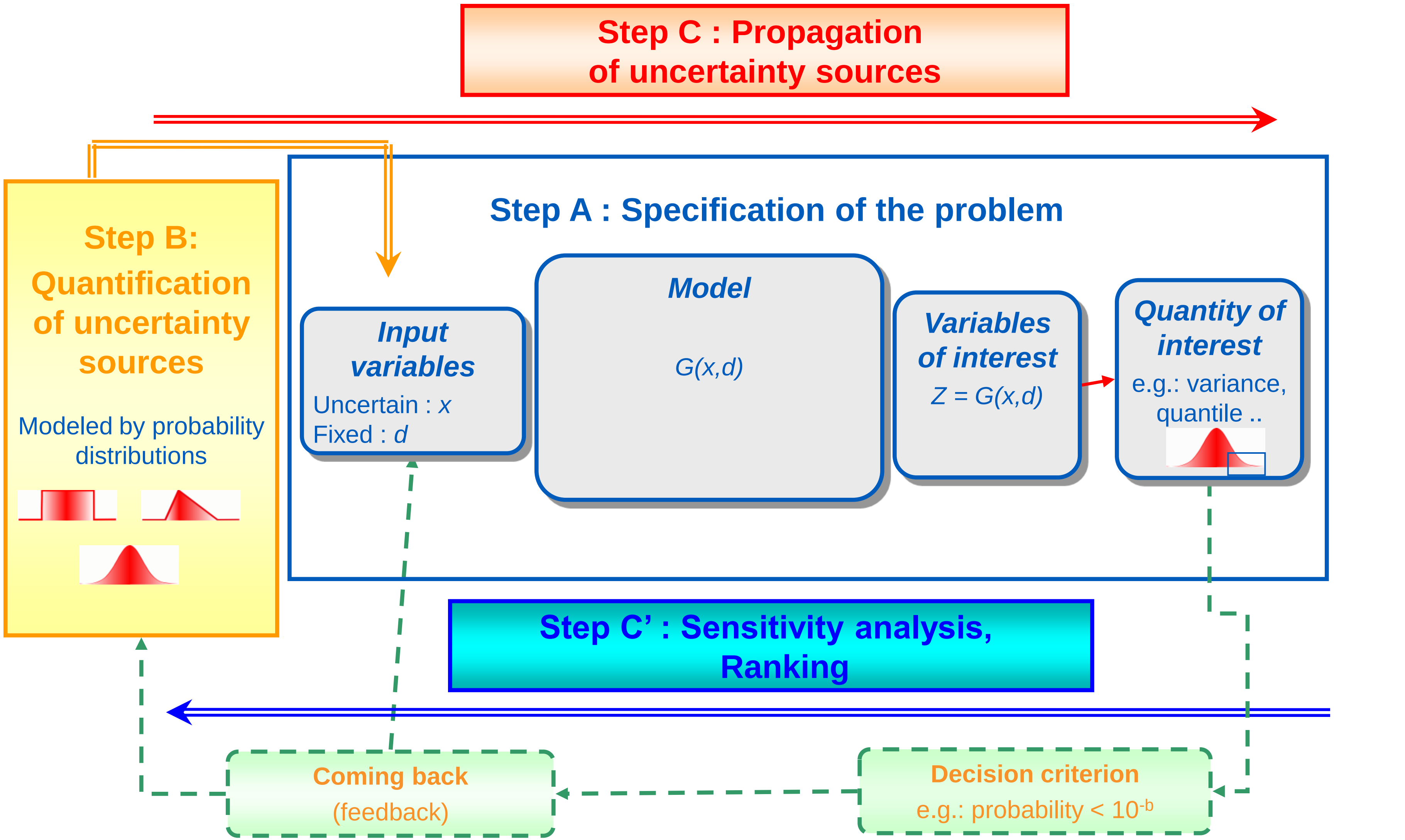
March 31, 2014



www.openturns.org

Contents

Uncertainty management - the global methodology



Propagating uncertainties: a tricky problem

Uncertainty analysis leads to the estimation of some particular features of the output *Y*

- An important example: the *probability of failure*
  - A system is in safe operating condition if *Y* is above (or below) a given "safety" threshold. Let us say if *Z*>0.

$$\mathcal{D}_f = \{x \in \mathcal{X} : G(x) \leq 0\}$$
$$p_f = \int_{\mathcal{D}_f} f(x)dx = \int_{\mathcal{X}} I_{\mathcal{D}_f}(x)\ f(x)dx = \mathbb{E}\left[I_{\mathcal{D}_f}(X)\right]$$

In practical application the order of magnitude of p_f could be 10^{-3} down to 10^{-7}

Most of time, sampling-based

- Brute force Monte Carlo methods are inappropriate: slow convergence leads to an unaffordable computational burden

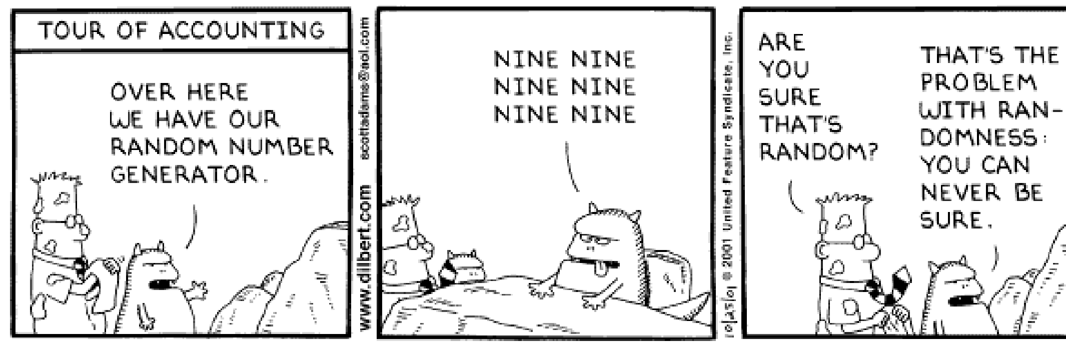
Different (and complementary) strategies:

- Adopting a sampling strategy which reduces the variance of the estimator of the desired quantity
- Building a surrogate model ("meta-model") the computational cost of which is null, then propagating uncertainty through the meta-model
  - The computational budget is here used to build the meta-model, following a particular DOE

Propagating uncertainties: take-home messages

- Main issue in practice with complex models: the computational burden!

- Avoid DIY solutions!



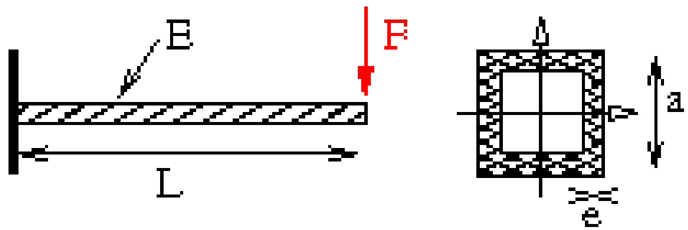
- You may need:
  - Appropriate math. methods (advanced sampling, meta-modeling)
  - Appropriate software tools for:
    - Effectively linking the deterministic model G(*X*) and the probabilistic model of *X*
    - Performing distributing computations (High Performance Computing)

The “Uncertainty Analysis Community”

- Several collaborative frameworks at the interface industry-academics
  - multi-partners projects, funded by public powers
  - Industrial partnerships for developing Open Source Software (e.g. OpenTURNS)
  - Research consortiums
  - Working groups within scientific societies

It’s up to you with OpenTURNS and Uranie

- Train yourself on a simple example:
  - deviation of a beam on which we apply a load when its other extremity is fixed in a walll
  - we model uncertainties on the parameters E, F, L, I
  - we apply some of the algorithms of the Methodology using OpenTURNS or Uranie



$$Y = \frac{FL^3}{3EI}$$

- statistical fitting
- dependency modeling
- mean and variance estimation
- sensitivity analysis
- rare event probability estimation using Monte Carlo and FORM estimators
- ...

Consider a diving board in a swimming pool where many kids are jumping....

What about the deviation of the diving board ... and the safety of people just under??

Thank you for your attention