Uncertainty quantification in the simulation of complex systems

Uncertainty quantification in the simulation of complex systems

Prof. Dr. Bruno Sudret

1st International Conference on Infrastructure Resilience

February 15th, 2018

A complex engineering system

Cruas nuclear power plant **(Source: EDF R&D)**

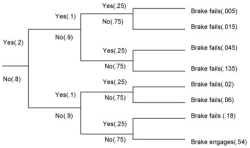
Quantitative risk analysis

A three-step general approach

Kaplan & Garrick, 1981

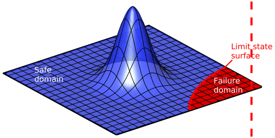
- What can go wrong?

  Scenarios of undesired events



- **How likely are the events?**

  Probabilities of failure



- What are the consequences?

  human / environmental / financial impact

Computational models for risk analysis

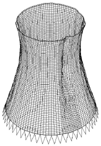
To assess the reliability of systems, computational models are used to predict the performance under different operating conditions

A computational model combines:

- A mathematical description of the physical phenomena (governing equations), *e.g.* mechanics, electromagnetism, fluid dynamics, etc.

$$\begin{aligned} \nabla \cdot \mathbf{D} &= \rho \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \end{aligned}$$

- Discretization techniques which transform continuous equations into linear algebra problems



- Algorithms to solve the discretized equations

Computational models: the abstract viewpoint

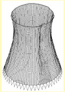
A computational model may be seen as a black box program that computes quantities of interest (QoI) as a function of input parameters

Vector of input parameters $\boldsymbol{x} \in \mathbb{R}^M$

Computational model $\mathcal{M}$

Model response $\boldsymbol{y} = \mathcal{M}(\boldsymbol{x}) \in \mathbb{R}^N$

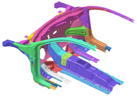
- Geometry
- Material properties
- Loading



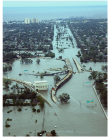
- Analytical formula
- Finite element model
- Comput. workflow

- Displacements
- Strains, stresses
- Temperature, etc.

Real world is uncertain



- Differences between the designed and the real system:
  - Dimensions (tolerances in manufacturing)
  - Material properties (*e.g.* variability of the stiffness or resistance)
- Unforecast exposures: exceptional service loads, natural hazards (earthquakes, floods, landslides), climate loads (hurricanes, snow storms, etc.), accidental human actions (explosions, fire, etc.)



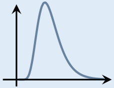
Global framework for uncertainty quantification

Step B
Quantification of sources of uncertainty

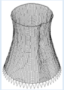
Step A
Model(s) of the system
Assessment criteria

Step C
Uncertainty propagation

Random variables



Computational model



Moments
Probability of failure
Response PDF

Step C'
Sensitivity analysis

B. Sudret, *Uncertainty propagation and sensitivity analysis in mechanical models – contributions to structural reliability and stochastic spectral methods*
Habilitation à diriger des recherches, Université Blaise Pascal (2007)

Step A: Computational models and criteria

Computational model

Vector of input parameters $\boldsymbol{x} \in \mathbb{R}^M$ — Computational model $\mathcal{M}$ — Quantities of interest $\boldsymbol{y} = \mathcal{M}(\boldsymbol{x}) \in \mathbb{R}^N$

Performance criterion

Failure $\Leftrightarrow \; g(\boldsymbol{y}, \boldsymbol{\theta}) \leq 0$ *e.g.* when $y \geq y_{adm}$

Example

In a nuclear reactor, the maximal temperature of fuel rods shall be smaller than the melting point:

$$g(\boldsymbol{X}) = T_{melt} - \max_{\text{rods } i} T_i^{comp.}(\boldsymbol{X})$$

Step B: Quantification of the sources of uncertainty

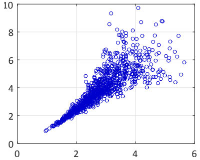
Goal: represent the uncertain parameters based on the *available data and information*

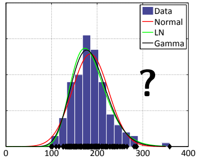
Probabilistic model $f_{\boldsymbol{X}}$

Experimental data is available

- What is the distribution of each parameter ?
- What is the dependence structure ?

Copula theory

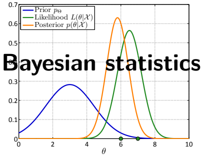




No data is available: expert judgment

- Engineering knowledge (*e.g.* reasonable bounds and uniform distributions)
- Statistical arguments and literature (*e.g.* extreme value distributions for climatic events)

Scarce data + expert information

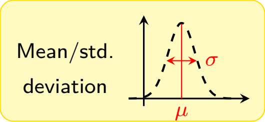


Step C: uncertainty propagation

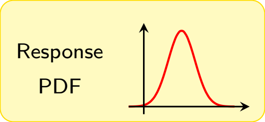
Goal: estimate the uncertainty / variability of the quantities of interest (QoI) $Y = \mathcal{M}(\boldsymbol{X})$ due to the input uncertainty $f_{\boldsymbol{X}}$

- Output statistics, *i.e.* mean, standard deviation, etc.

$$\mu_Y = \mathbb{E}_{\boldsymbol{X}}\left[\mathcal{M}(\boldsymbol{X})\right]$$
$$\sigma_Y^2 = \mathbb{E}_{\boldsymbol{X}}\left[\left(\mathcal{M}(\boldsymbol{X}) - \mu_Y\right)^2\right]$$

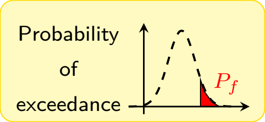


- Distribution of the QoI



- Probability of exceeding an admissible threshold y_{adm}

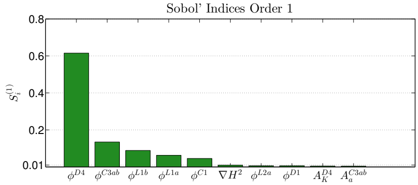
$$P_f = \mathbb{P}\left(Y \geq y_{adm}\right)$$



Step C': sensitivity analysis

Goal: Determine what are the input parameters (or combinations thereof) whose uncertainty explains the variability of the quantities of interest

- Screening: detect input parameters whose uncertainty has no impact on the output variability
- Feature setting: detect input parameters which allows one to best decrease the output variability when set to a deterministic value
- Exploration: detect interactions between parameters, *i.e.* joint effects not detected when varying parameters one-at-a-time



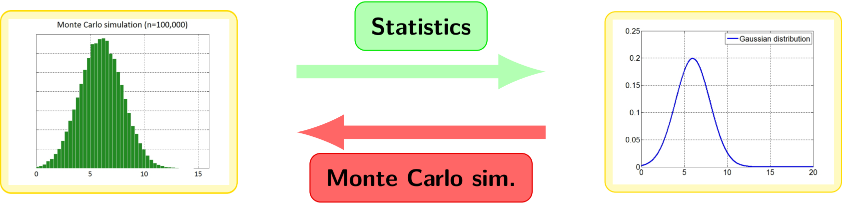
Variance decomposition: Sobol' indices

Outline

Monte Carlo simulation

Principle: reproduce numerically the variability of the input parameters using random numbers

- A sample set $\mathcal{X} = \{\boldsymbol{x}_1, \, \dots \, , \boldsymbol{x}_n\}$ is drawn according to the input distribution $f_{\boldsymbol{X}}$



- For each sample the quantity of interest (resp. performance criterion) is evaluated, say $\mathcal{Y} = \{\mathcal{M}(\boldsymbol{x}_1), \, \dots \, , \mathcal{M}(\boldsymbol{x}_n)\}$

- The set of quantities of interest is used for moments-, distribution- or reliability analysis

Advantages/Drawbacks of Monte Carlo simulation

Advantages

- Universal method: only rely upon simulating random numbers ("sampling") and running repeatedly the computational model
- Suited to High Performance Computing: "embarrassingly parallel"
- Sound statistical foundations: convergence when $N_{MCS} \to \infty$

Drawbacks

- Statistical uncertainty: results are not exactly reproducible when a new analysis is carried out (handled through confidence intervals)
- Low efficiency for reliability analysis

Example

Suppose $P_f = 10^{-3}$

- At least 1,000 samples are needed in order to observe one single failure (in the mean!)
- About 100 times more (*i.e.* 100,000 samples) are required to have a $\pm 10\%$ accuracy

Outline

Surrogate models for uncertainty quantification

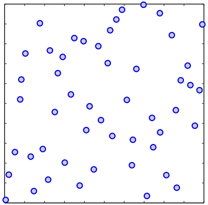
A surrogate model $\tilde{\mathcal{M}}$ is an approximation of the original computational model $\mathcal{M}$ with the following features:

- It is built from a limited set of runs of the original model $\mathcal{M}$ called the experimental design $\mathcal{X} = \{\boldsymbol{x}_i, \, i = 1, \, \dots \, , m\}$
- It assumes some regularity of the model $\mathcal{M}$ and some general functional shape

Name	Shape	Parameters
Polynomial chaos expansions	$\tilde{\mathcal{M}}(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} a_{\boldsymbol{\alpha}} \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x})$	$\boldsymbol{a}_{\boldsymbol{\alpha}}$
Low-rank tensor approximations	$\tilde{\mathcal{M}}(\boldsymbol{x}) = \sum_{l=1}^{R} b_l \left(\prod_{i=1}^{M} v_l^{(i)}(x_i) \right)$	$b_l, \, z_{k,l}^{(i)}$
Kriging (a.k.a Gaussian process modelling)	$\tilde{\mathcal{M}}(\boldsymbol{x}) = \boldsymbol{\beta}^{\mathsf{T}} \cdot \boldsymbol{f}(\boldsymbol{x}) + Z(\boldsymbol{x}, \omega)$	$\boldsymbol{\beta} \, , \, \sigma_Z^2 \, , \, \boldsymbol{\theta}$
Support vector machines	$\tilde{\mathcal{M}}(\boldsymbol{x}) = \sum_{i=1}^{m} a_i \, K(\boldsymbol{x}_i, \boldsymbol{x}) + b$	$\boldsymbol{a} \, , \, b$

- It is fast to evaluate

Ingredients for building a surrogate model



- Select an experimental design $\mathcal{X}$ that covers at best the domain of input parameters: Latin Hypercube Sampling, low-discrepancy sequences
- Run the computational model $\mathcal{M}$ onto $\mathcal{X}$ exactly as in Monte Carlo simulation
- Smartly post-process the data $\{\mathcal{X}, \mathcal{M}(\mathcal{X})\}$ through a learning algorithm

Name	**Learning method**
Polynomial chaos expansions	sparse grids, regression, LAR
Low-rank tensor approximations	alternate least squares
Kriging (a.k.a Gaussian process modelling)	maximum likelihood, Bayesian inference
Support vector machines	quadratic programming

Advantages/Drawbacks of surrogate models

Usage

$$\mathcal{M}(\boldsymbol{x}) \quad \approx \quad \tilde{\mathcal{M}}(\boldsymbol{x})$$

hours per run — seconds for 10^6 runs

Advantages

- Non intrusive methods: based on runs of the computational model, exactly as in Monte Carlo simulation
- Suited to High Performance Computing: "embarrassingly parallel"
- Similarities with big data analysis (offline interaction with computational models)
- Efficiency: 2-3 orders of magnitude less runs compared to Monte Carlo

Drawbacks

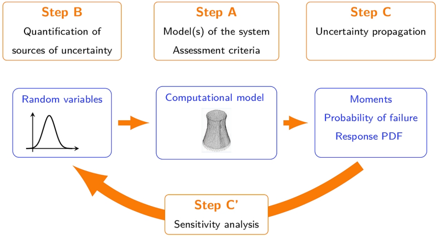
- Need for rigorous validation of the surrogate model
- Advanced mathematical background: not easy to communicate to practitioners

UQLab

The Framework for Uncertainty Quantification

OVERVIEW | FEATURES | DOCUMENTATION | DOWNLOAD/INSTALL | ABOUT

"Make uncertainty quantification available for anybody, in any field of applied science and engineering"



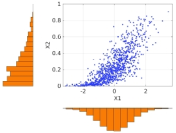
- MATLAB®-based Uncertainty Quantification framework
- State-of-the art, highly optimized open source algorithms
- Steep learning curve for beginners
- Modular structure, easy to extend
- Exhaustive documentation

http://www.uqlab.com

UQLab features

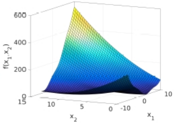
Probabilistic Input Modelling

- Common marginals
- Support for user-defined marginals
- Support for bounds on all distributions (including user-defined)
- Gaussian copula



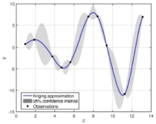
Modelling facilities

- Simple text strings
- MATLAB m-files
- MATLAB handles
- Simple API to produce wrappers to commercial/external solvers



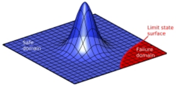
Advanced metamodelling

- Sparse degree-adaptive Polynomial Chaos Expansions
- Gaussian process modelling (Kriging)
- Polynomial-Chaos Kriging
- Low-rank tensor approximations



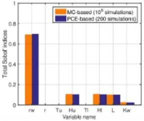
Reliability Analysis (Rare Event Estimation)

- FORM/SORM approximation
- Monte Carlo Simulation (MCS)
- Importance Sampling
- Subset Simulation
- Adaptive Kriging (AK-MCS)



Sensitivity Analysis

- Correlation-based indices
- Standard Regression Coefficients
- Cotter measure
- Morris indices
- Sampling-based Sobol' indices
- PCE-based Sobol' indices



Upcoming features

- UQLink: easily connect UQLab to external modelling software
- Bayesian model calibration/inversion toolbox
- Random fields discretization and sampling toolbox
- Support vector machines for regression and classification
- Reliability-based design optimization (RBDO)
- Advanced dependence modelling and inference with vine copulas

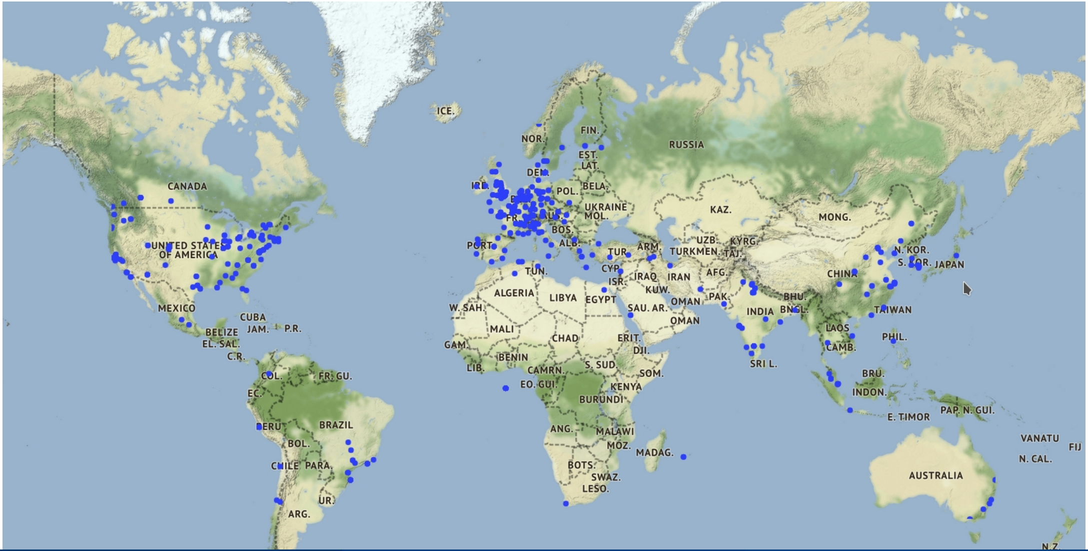
Facts and figures

http://www.uqlab.com

- Development since 2013, first beta version V0.9 on July 1st, 2015
- Release of V1.0 on April 28, 2017
- **ETH license**:
  - + free of charge for academia
  - + yearly fee for non-academic usage
- **1180 users** (680 active) in **57 countries**

Country	**# licences**
United States	225
France	142
Switzerland	117
China	92
Germany	74
United Kingdom	69
Italy	43
India	35
Canada	32
Belgium	28

UQLab users

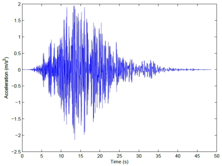

ICE.
NOR.
FIN.
EST.
LAT.
DEN.
IRL.
POL.
BELA.
UKRAINE
MOL.
FR.
BOS.
ALB.
PORT.
TUN.
RUSSIA
KAZ.
MONG.
UZB.
KYRG.
TURKMEN.
TAJ.
TUR.
ARM.
CYP.
ISR.
IRAQ
IRAN
AFG.
KUW.
PAK.
OMAN
SAU. AR.
CHINA
N. KOR.
S. KOR.
JAPAN
BHU.
INDIA
BNGL.
TAIWAN
LAOS
CAMB.
PHIL.
SRI L.
BRU.
INDON.
E. TIMOR
PAP. N. GUI.
VANATU
FIJ
N. CAL.
AUSTRALIA
N.Z.
CANADA
UNITED STATES
OF AMERICA
MEXICO
CUBA
JAM.
P.R.
BELIZE
EL. SAL.
C.R.
COL.
FR. GU.
EC.
PERU
BRAZIL
BOL.
CHILE
PARA.
UR.
ARG.
ALGERIA
LIBYA
EGYPT
W. SAH.
MALI
CHAD
ERIT.
DJI.
GAM.
BENIN
LIB.
CAMRN.
S. SUD.
SOM.
EQ. GUI.
KENYA
BURUNDI
ANG.
MALAWI
MOZ.
MADAG.
BOTS.
SWAZ.
LESO.

UQLab users


ETH Zurich
University of Illinois at Urbana-Champaign
Universiteit Gent
Technical University of Denmark
Université de Nantes
Technische Universität München
Télécom ParisTech
University of Liverpool
Cranfield University
Bauhaus Universität Weimar
San Francisco State University
Institut National des Sciences Appliquées de Rouen
University of Tehran
Virginia Polytechnic Institute and State University
National University of Defense Technology
Universidad Nacional de Colombia
University of Sheffield
Stanford University
Pontificia Universidad Católica del Perú
Institut Fresnel
University of Zurich
University of Colorado at Boulder
Beijing Institute of Applied Physics and Computational Mathematics
University of Maryland, College Park
Université d'Angers
Sapienza - Università di Roma
Rheinisch Westfälische Technische Hochschule Aachen
Hunan University
Texas A&M University - College Station
King Abdullah University of Science and Technology
University of Michigan - Ann Arbor
Indian Institute of Technology, Madras
Huazhong University of Science and Technology
University of Florida
Université de Strasbourg
University of Minnesota
Yeungnam University
University of Crete
Université de Lorraine
University of Surrey
Iowa State University
Ohio State University, Columbus
Chalmers University of Technology
Johns Hopkins University
Tohoku University
Technische Universität Darmstadt
Pontifícia Universidade Católica do Rio de Janeiro
Katholieke Universiteit Leuven
National University of Singapore
Beijing Jiaotong University
Bocconi University
CentraleSupélec
Sistan & Balouchestan University
Shanghai Jiao Tong University
University of Science and Technology of China
Arts et Métiers ParisTech
Ecole Supérieure d'Electricité
Tongji University
University of Leeds
Indian Institute of Technology, Delhi
San Mateo County Community College District
Beihang University
Zhejiang University
Aalborg University
University of Notre Dame
Ruhr-Universität Bochum
European Commission - Joint Research Centre
Max Planck Institute - Magdeburg
Cornell University
Central South University
University of Exeter
Technische Universität Wien
Polytechnic Institute of Turin
Harbin Institute of Technology
Indian Institute of Technology, Bombay
Norwegian University of Science and Technology
Swansea University
Laboratoire 3SR
Monash University
French Geological Survey
Tsinghua University
University of Twente
Universidad Politécnica de Madrid
Kunsan National University
Polytechnic Institute of Milan
Technical University of Budapest
University of New South Wales
Northwest Polytechnical University Xi'an
University of Ottawa
North Carolina State University
Lawrence Livermore National Laboratory
Universidade Federal de Santa Catarina
Indian Institute of Science
Singapore University of Technology and Design
Université Libre de Bruxelles
EPFL - EPF Lausanne
University of Bristol
University of Utah
INRIA
ONERA - Office National d'Etudes et de Recherches Aérospatiales
IRSTEA
University of Florence
Aalto University
Université de Technologie de Compiègne
Southeast University
State University of New York at Buffalo
Universität Stuttgart
Nanyang Technological University
Universidade Federal do Rio de Janeiro
Technische Universität Carolo-Wilhelmina Braunschweig
Delft University of Technology
University of Texas at Austin
Carleton University
National Technical University of Athens
Sharif University of Technology
Ecole Normale Supérieure de Cachan
Technische Universität Braunschweig
Université Paris-Est-Marne-la-Vallée
Nanjing University of Aeronautics and Astronautics
Centrum Wiskunde & Informatica Amsterdam
Universidad Autónoma del Estado de México
SIGMA Clermont
Eindhoven University of Technology
McMaster University
University of Connecticut
Universidad de Salamanca
Imperial College London
University of Warwick
Purdue University
Wuhan University
University of California, Berkeley
Technische Universität Chemnitz
Universität der Bundeswehr München
University of Delaware
Indian Institute of Technology, Roorkee
University of Calgary
Rice University
Massachusetts Institute of Technology
Universidade de Lisboa
University of California, Davis
Royal Melbourne Institute of Technology
Vanderbilt University
SP Technical Research Institute of Sweden
University of Toronto
Third University of Rome
Universidad de Los Andes
Université Claude Bernard (Lyon I)
University of Oxford
Universidade do Porto
Technische Universität Graz

Outline

Risk in earthquake engineering



Question

What is the probability of collapse of a building as a function of the "intensity" of a potential earthquake?

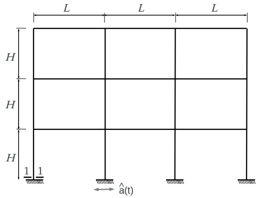
Uncertainties

- Properties of the structure (material strength, stiffness of the connections, etc.)
- Earthquake magnitude, duration, peak ground acceleration

Non linear transient finite element analysis of the structure
for different synthetic earthquakes

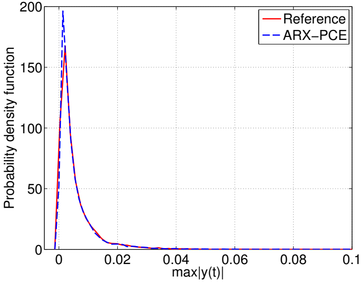
Risk in earthquake engineering

C.V. Mai, PhD Thesis (2015)

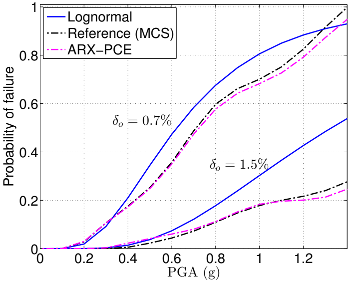


Computation of fragility curves: $F(PGA) = \mathbb{P}\left(\text{Drift} > \delta_0 | PGA\right)$

- Uncertainty in the steel Young's modulus and strength
- 10,000 synthetic earthquakes as a reference, 300 only for the surrogate model

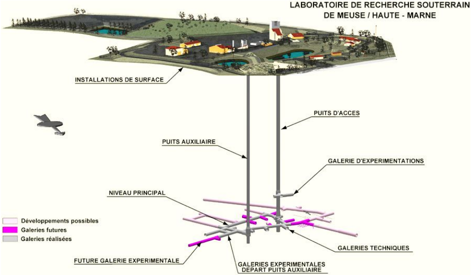


Distribution of peak drift



Fragility curves

Risk in nuclear waste storage



Source: http://www.futura-sciences.com/

Source: http://lexpansion.lexpress.fr/

- When assessing a nuclear waste repository, the Mean Lifetime Expectancy MLE(x) is the time required for a molecule of water at point x to get out of the boundaries of the system
- Computational models have numerous input parameters (in each geological layer) that are difficult to measure, and that show scattering

Geological model

Joint work with University of Neuchâtel

Deman, Konakli, Sudret, Kerrou, Perrochet & Benabderrahmane, Reliab. Eng. Sys. Safety (2016)

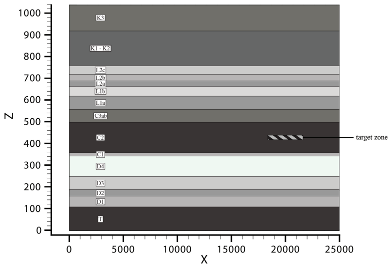
- Two-dimensional idealized model of the Paris Basin (25 km long / 1,040 m depth) with 5×5 m mesh (10^6 elements)
- Steady-state flow simulation with Dirichlet boundary conditions:

$$\nabla \cdot (\mathbf{K} \cdot \nabla H) = 0$$

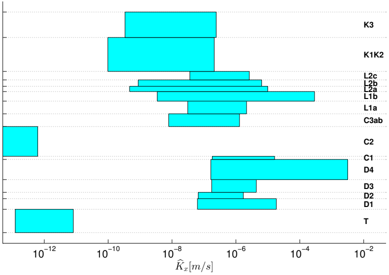
- 15 homogeneous layers with uncertainties in:
  - Porosity (resp. hydraulic conductivity)
  - Anisotropy of the layer properties (inc. dispersivity)
  - Boundary conditions (hydraulic gradients)

78 input parameters

Sensitivity analysis



Geometry of the layers



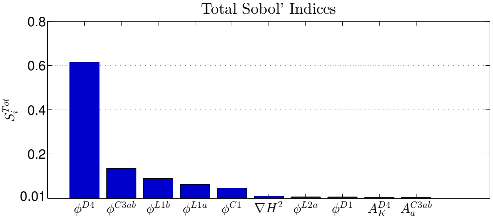
Conductivity of the layers

Question

What are the parameters (out of 78) whose uncertainty drives the uncertainty of the prediction of the mean life-time expectancy?

Sensitivity analysis: results

Technique: Sobol'indices computed from polynomial chaos expansions



Parameter	$\sum_j S_j$
ϕ (resp. K_x)	0.8664
A_K	0.0088
θ	0.0029
α_L	0.0076
A_α	0.0000
∇H	0.0057

Conclusions

- Only 200 model runs allow one to detect the 10 important parameters out of 78
- Uncertainty in the porosity/conductivity of 5 layers explain 86% of the variability
- Small interactions between parameters detected

Conclusions

- Uncertainty quantification has become a hot topic in all fields of applied sciences and engineering
- It plays a key role in risk assessment of complex systems described by advanced computational models
- There is a world beyond Monte Carlo simulation for risk analysis
- Surrogate modelling (*e.g.* polynomial chaos expansions, Kriging, etc.) is becoming a standard technique that decreases the computational costs by 2-3 orders of magnitude for the same accuracy
- UQ techniques can be made generic, versatile and field-independent, see ongoing applications in:
  - Macroeconomics — Risk Center / ETH D-MTEC
  - Urban water management — EAWAG
  - Crashworthiness of vehicles — Peugeot (France)
  - Astrophysics (!) — University of Zurich
  - ...

Questions ?

Chair of Risk, Safety & Uncertainty Quantification

www.rsuq.ethz.ch

The Uncertainty Quantification Laboratory

www.uqlab.com

Thank you very much for your attention !