EADS Innovation Works - GDR CNRS MASCOT-NUM

Uncertainty Management and High Performance Computing : what is at stake from an applicative point of view ?

Fabien MANGEANT, fabien.mangeant@eads.net

March 28, 2013

Outline

Acknowledgements

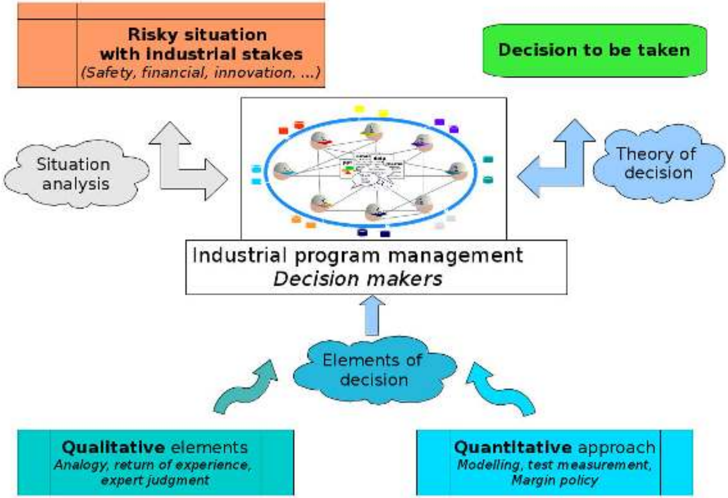
This presentation is the result of many discussions in the frame of the **GDR CNRS MASCOT NUM** (*Anestis Antoniadis, Jean-Claude Fort, Fabrice Gamboa, Josselin Garnier, Erick Herbin, Bertrand Iooss, Antony Nouy, Clementine Prieur, Luc Pronzato, Bruno Sudret...*), the **ImdR and EsREDA groups** (*Nicolas Devictor, Nicolas Fischer, Etienne de Rocquigny, ...*), the **Open TURNS consortium** (*Anne Dutfoy, Alberto Pasanisi, Thierry Yalamas, ...*), the OPUS project (Christophe Prudhomme [25], ...).

A special thanks to my **EADS IW** colleagues (*Pierre Benjamin, Eric Duceau, Vincent Feuillard, Regis Lebrun, Nabil Rachdi, Jayant Sen Gupta, Benoit Lize, Guillaume Sylvand*) for all the fruitful and stimulating projects and discussions on the subject and its applications in aeronautics and aerospace !

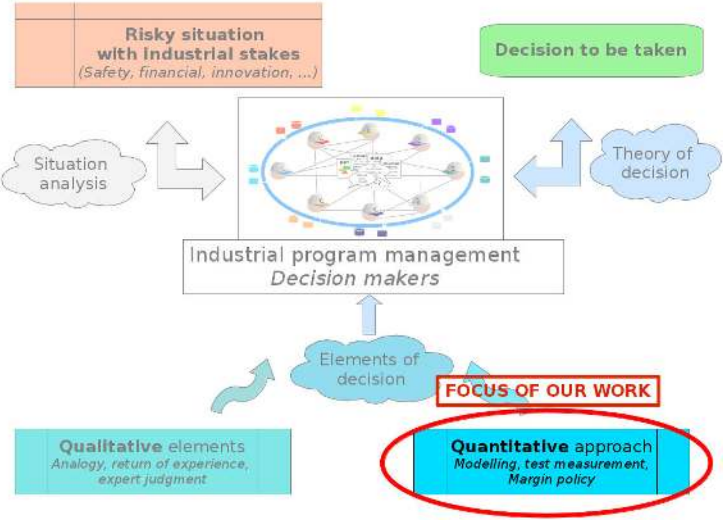
Outline

1. **Motivations**
   - General context
   - Uncertainty Management
   - Engineering challenges
   - Types of uncertainty
   - Uncertainty management
2. **HPC for UM in an industrial environment**
   - HPC for uncertainty management in preliminary design
   - HPC for uncertainty management in virtual testing
   - Some technological realizations
3. **Summary and perspectives**
   - Challenges associated to the deployment of such approaches
4. **Bibliography**

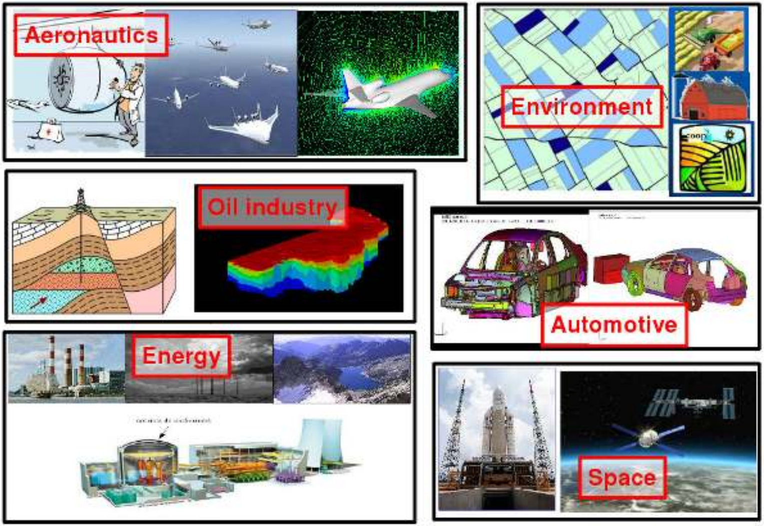
Uncertainty analysis in a decision process



Uncertainty analysis in a decision process



A multi sector concern

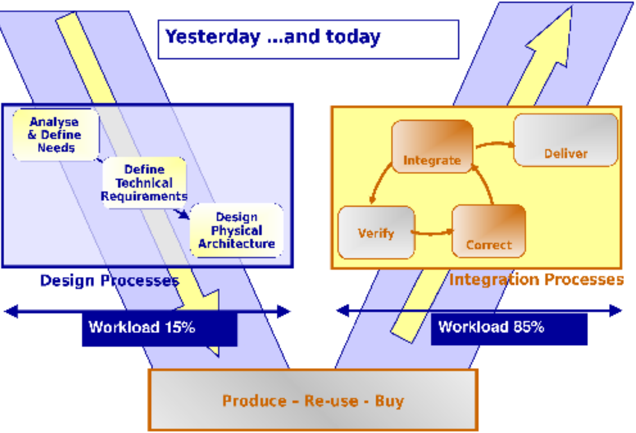


New opportunities to improve procedures and practices around uncertainty management

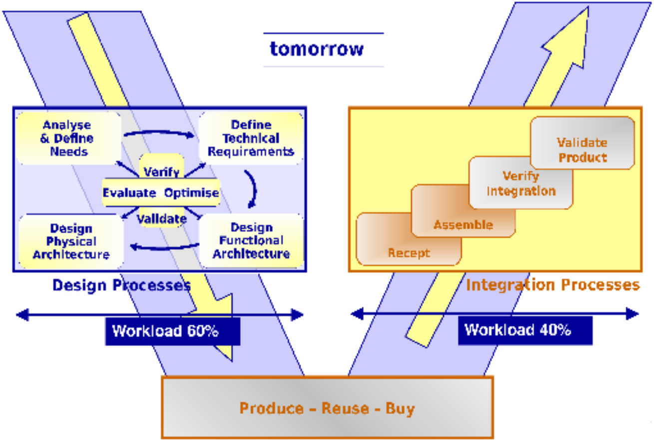
1. Recent conceptual reformulation: **shift from "failure-driven risk approach" to "option-exploring approach"** in uncertainty management.
2. Recognition that **the performance of an engineered system has to be taken into account in its larger commercial and political environment**:
   - To bullet-proof design against technical and human failure
   - To enable the system to evolve to new circumstances
3. **New advances in information technology** (development of models, acquisition of computer, etc) make it possible to conceive of a much more coherent uncertainty management approach.

See the references linked to systems engineering and risk management literature: Apostolakis [2], Aven [3], Bedford [4], Garnger [11], Knight [14], Neufville [7, 8], de Weck [32]

HPC for UM in an industrial environment



HPC for UM in an industrial environment



EADS

UM environment

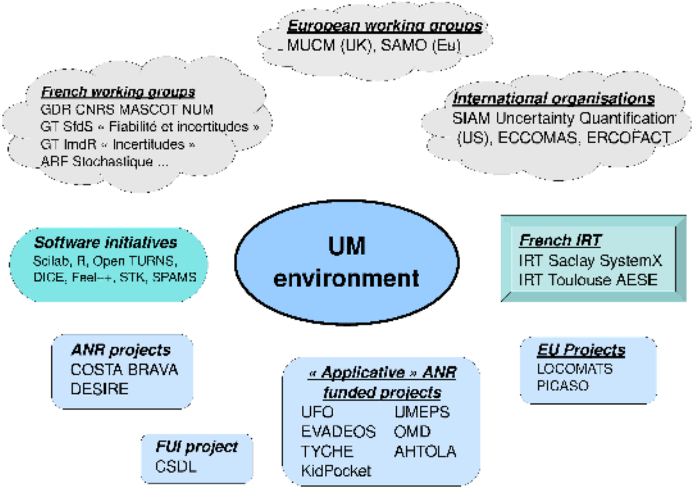
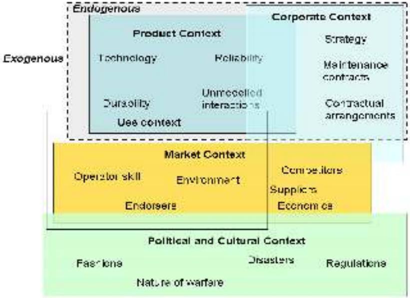
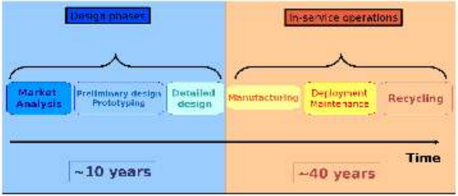


Figure: On-going initiatives in UM that I know... See for example MASCOT NUM [19] or MUCM [22]

EADS

An architect view





Performances

Aerodynamic: Drag,

Mass: Maximum Weight,

Acoustics: Perceived Noise Level,

Energy: Maximum Electric Power,

Propulsion: Specific Fuel Consumption...

$$\Downarrow$$

$$\mathbf{y} = (y_1, \ldots, y_Q)$$

A naive presentation of the engineering challenge

Description of the situation

- A target $\mathcal{T}$ is given to the variable $\mathbf{y}^*$. This target can evolve during the time of the design.
- These performances are uncontrolled for many reasons (lack of knowledge, variability, approximation, dependency, ...).
- The amount of available information $\mathcal{I}$ for each variable y_i^* evolves during the time of the design (either over the knowledge of the input variables, parameters, mesurements, availability of numerical models).
- At a given time of the design, these technical performances must be estimated with a level of confidence.

A naive presentation of the engineering challenge

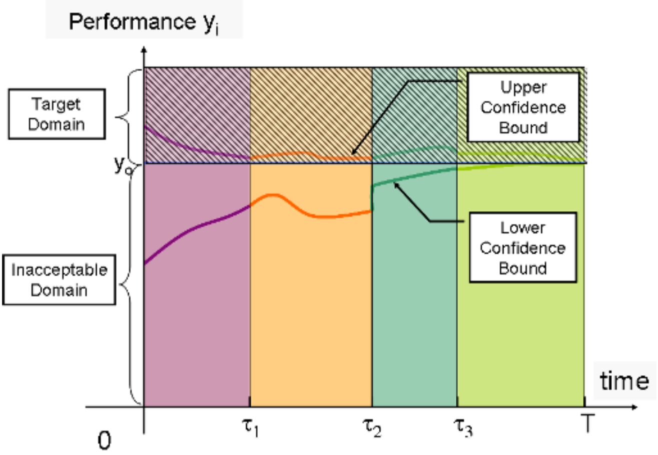


Figure: Evolution of a performance during the design phase

A naive presentation of the engineering challenge

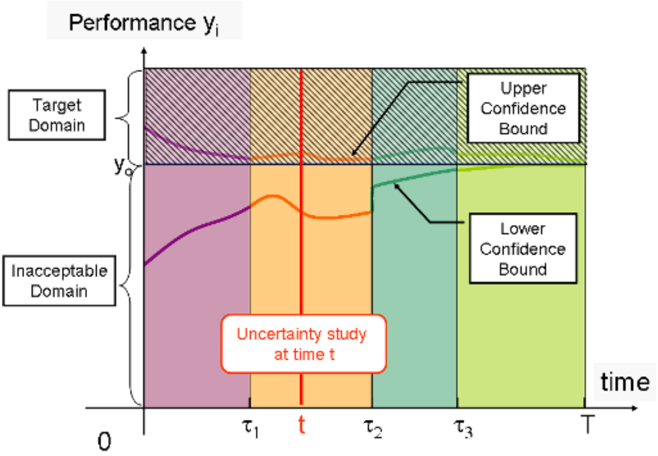
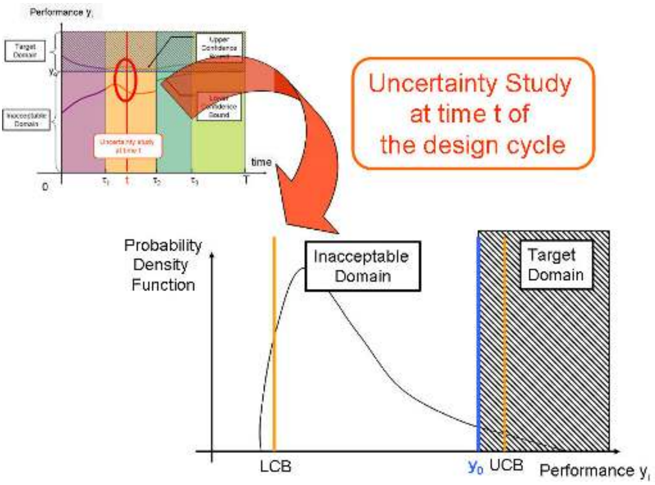


Figure: An uncertainty study at a given time of the design

A naive presentation of the engineering challenge



A naive presentation of the engineering challenge

Objectives in a mathematical framework

In a probabilistic framework, two main goals can be identified:

1. To control the stochastic behaviour of the performances $\mathbf{y}^*$ to reach the initial or adapted target $\mathcal{T}$.
2. To estimate on-demand some measures of risks $\rho(\mathbf{Y}^*)$ during the time of the design.

A naive presentation of the engineering challenge

Objectives in a mathematical framework

In a probabilistic framework, two main goals can be identified:

1. To control the stochastic behaviour of the performances $\mathbf{Y}^*$ to reach the initial or adapted target $\mathcal{T}$.
2. **To estimate on-demand one or several measures of risks $(\rho_i(\mathbf{Y}^*))_{i=1,\cdots,d}$ during the time of the design. –> This is a new discipline for engineers and where we focus our current efforts!**

What kind of information do we manipulate?

Elements of information

- A **reference database** $(\mathbf{Y}_1^*, \cdots, \mathbf{Y}_n^*)$ or $((\mathbf{X}_1^*, \mathbf{Y}_1^*), \cdots, (\mathbf{X}_n^*, \mathbf{Y}_n^*))$ that is enriched during the design cycle.
- A **set of risk measures** $(\rho_1(\mathbf{Y}^*), \cdots, \rho_d(\mathbf{Y}^*))$ built upon $\mathbf{Y}^*$ to be estimated during the design cycle.
- A **panoply of numerical models** $\mathcal{H} = \{h_1, \cdots, h_D\}$ that is enriched during the design cycle.
- A **quantification of the uncertainties attached to the inputs** of the numerical models represented **by a statistical law** $\mathbb{P}_{\mathbf{X}}$ that is enriched during the design cycle
- A **definition of the target** $\mathcal{T}$ and its associated level of confidence α to be reached that is enriched during the design cycle.
- A **global computational budget** $\mathcal{B}$ that can be allocated at different times of the design cycle.

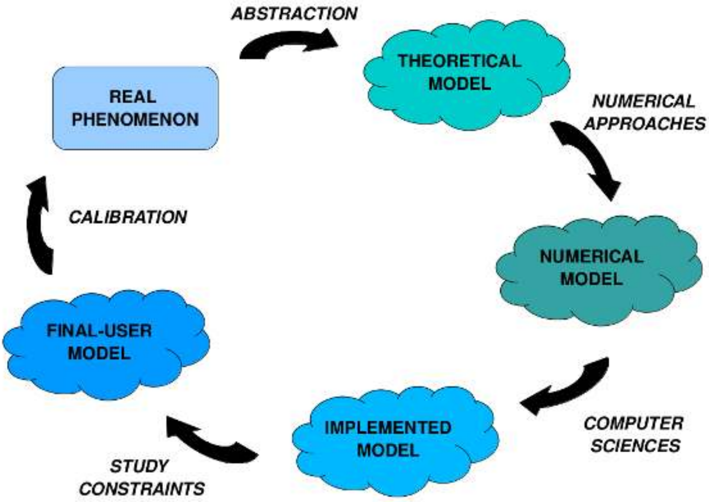


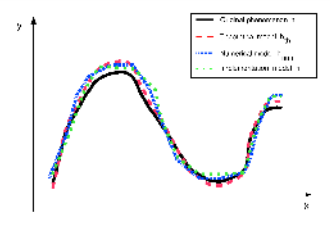
Figure: Modelling loop

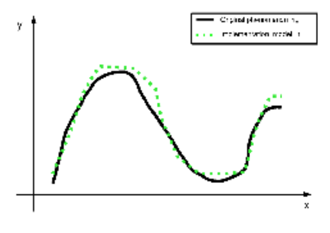
What is the nature of uncertainty in this context?

"Model" uncertainty

- **Reference model** h^*: Usually not accessible, expression of a natural or a complex technical object.
- **Theoretical model** h_{th}: Scientific expert activity (modelling activity, theoretical solution of a PDE system, ...), corresponding to the level of understanding and simplification of the problem.
- **Numerical model** h_{num}: Numerical solution of the theoretical model (effects of meshing, choice of a numerical scheme, truncature effects, ...)
- **Implementation model** h: Software implementation of the model on a given hardware architecture (computer accuracy, choice of coding rules, ...)

$$h^* \rightsquigarrow h$$





What is the nature of uncertainty in this context?

Parametric input uncertainty

- For a given numerical model $h : (\mathbf{x}, \theta) \in \mathcal{X} \times \Theta \mapsto \mathbf{y} = h(\mathbf{x}, \theta) \in \mathcal{Y}$, we consider **an uncertainty attached to the input variables X modelled by a statistical law** $\mathbb{P}^*_{\mathbf{X}}$.
- In practical contexts, it is often difficult to build $\mathbb{P}^*_{\mathbf{X}}$ due to scarsity of data, heterogeneous database, lack of information on the dependency, ... As a matter of fact, one has to work with an **approximate statistical law** $\mathbb{P}_{\mathbf{X}}$.

$$\mathbb{P}^*_{\mathbf{X}} \rightsquigarrow \mathbb{P}_{\mathbf{X}}$$

Computational budget $\mathcal{B}$

- In many situations, it is difficult to compute analytically the risk measures $\rho(h(\mathbf{X}, \theta))$. Numerical methods $\mathcal{M}(\mathcal{B}, \varepsilon, h(\mathbf{X}, \theta)$ (either stochastic or not) are required using a fixed computational budget $\mathcal{B}$ for a given accuracy ε

$$\rho(h(\mathbf{X}, \theta)) \rightsquigarrow \mathcal{M}(h(\mathbf{X}, \theta), \mathcal{B}, \varepsilon)$$

How to manage all the components of the error?

Recap of the errors

1. Building of the model: $\mathcal{N}_{\mathcal{S}}(h^*, h_{th})$
2. Numerical approximation: $\mathcal{N}_{\mathcal{N}}(h_{th}, h_{num})$
3. Hardware/Software implementation: $\mathcal{N}_{\mathcal{I}}(h_{num}, h)$
4. Model paramaters uncertainty: $\mathcal{N}_{\mathcal{Q}}(\mathbb{P}^*_{\mathbf{X}}, \mathbb{P}_{\mathbf{X}})$
5. Uncertainty propagation error: $\mathcal{N}_{\mathcal{P}}(\rho(h(\mathbf{X}, \theta)), \mathcal{M}(h(\mathbf{X}, \theta), \mathcal{B}, \varepsilon))$

Naive form of the total error

$$\begin{aligned}\Delta \leq \quad & \underbrace{\mathcal{N}_{\mathcal{S}}(h^*, h_{th})}_{\textit{Scientific Validation}} \\ & + \underbrace{\mathcal{N}_{\mathcal{N}}(h_{th}, h_{num})}_{\textit{Numerical Validation}} + \underbrace{\mathcal{N}_{\mathcal{I}}(\hat{h}, h)}_{\textit{Hardware/Software Validation}} \\ & + \underbrace{\mathcal{N}_{\mathcal{Q}}(\mathbb{P}^{\mathbf{X}}_{*}, \mathbb{P}^{\mathbf{X}})}_{\textit{Statistical Validation}} + \underbrace{\mathcal{N}_{\mathcal{P}}(\rho(Y), \hat{\rho}_{\mathcal{B}}(Y))}_{\textit{Propagation Validation}}\end{aligned}$$

How to manage some components of the error?

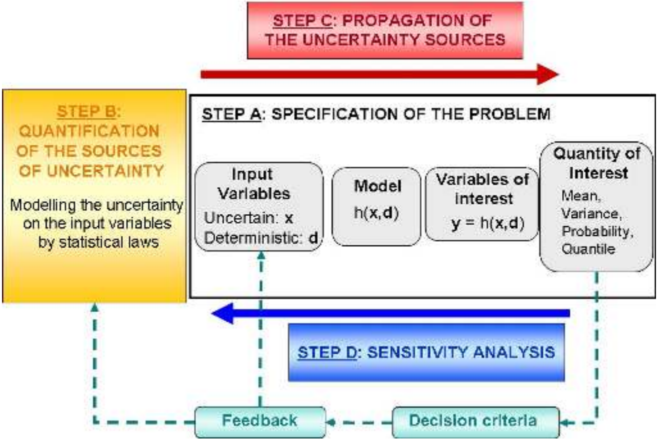


Figure: Four steps engineering process for uncertainty study (ref EsREDA [9], Open TURNS [23])

Outline

HPC for uncertainty management in preliminary design

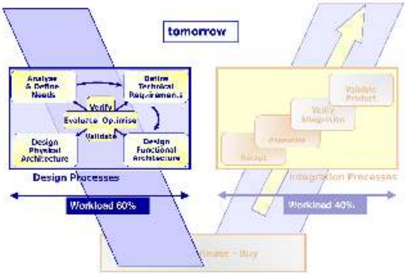


Figure: Modelling activity in early design

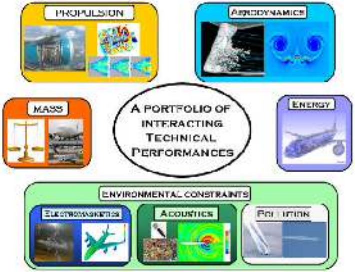


Figure: Portfolio of interacting technical performances

HPC for uncertainty management in preliminary design

Topic	HPC need	Context
Model complexity	+	Simplification of the modelling
Definition of the product	-	Missing data
Uncertainty	++	Large uncertainty on product definition, model definition, ...
Multi disciplinary coupling	++	Weak coupling but with numerous disciplines
Computational unit	++	Large set of numerical experiments
Memory required	+	

Table: HPC requirement in early design

HPC for uncertainty management in preliminary design

HPC constraints for preliminary design

- Small computational units (stand-alone PCs, small cluster)
- Large grid: distribution of these unitary units among the company and outside (extended enterprise)
- Easy interface of HPC capabilities for the end-user (who is rarely a specialist in numerical simulation)
- Securization of the data

HPC for uncertainty management in virtual testing

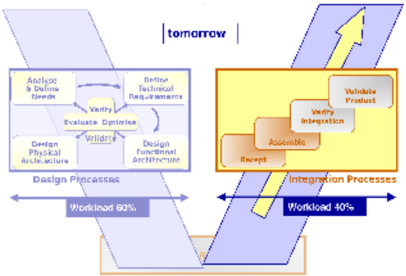


Figure: Modelling activity in virtual testing

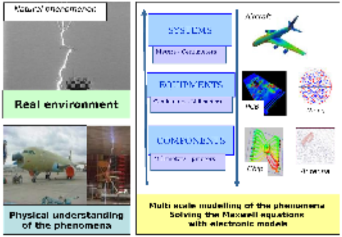


Figure: Multi scale modelling in electromagnetism

EADS

HPC for uncertainty management in virtual testing

Topic	HPC need	Context
Model complexity	++	High accuracy required
Definition of the product	++	Difficulty to transform the geometry into a simulation model
Uncertainty	+	Less uncertainty, high accuracy on criteria
Multi disciplinary coupling	++	Strong coupling between few disciplines
Computational unit	++	Large model experiments
Memory required	++	Large meshings with many unknowns

Table: HPC requirement in virtual testing

HPC for uncertainty management in virtual testing

HPC constraints for virtual testing

- Large computational units (big clusters, parallel machines)
- Grid or grid of clusters: synchronous resolution problem by various owners of the problem (extended enterprise)
- Securization of the data and models (either by numerical resolution or data transfer)

Some technological realizations

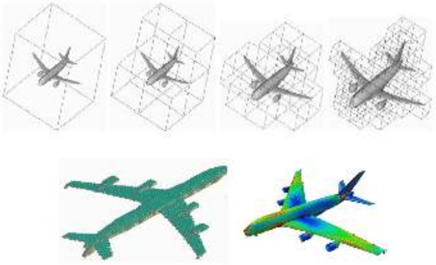


Figure: Multi scale modelling in electromagnetism

Evolution

- Since 1990 : Dense Linear algebra, Direct solver, up to 16 processor
- Since 2000 : Fast Algorithm (FMM), Iterative solver, up to 64 processors
- Today : Used in Buâs everyday, up to 500 processors

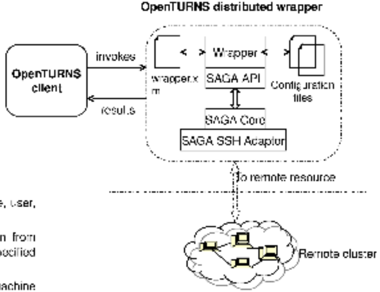
Some technological realizations

OpenTURNS Distributed Wrapper

What is it for ?

Supports execution of OpenTURNS based simulations in remote cluster distributing them in multiple nodes of the cluster by

- ✓ Transferring required input files to the remote machine (which contains the scientific code)
- ✓ Modifies input files based on the control factors of the experiment
- ✓ Invokes appropriate command to execute the model in the remote machine, and transfers output files from the remote execution machine



How it does ?

- ➢ Uses pre-defined configuration files for identifying remote resource, user, remote working directory and information about shared file system
- ➢ Creates a shell script (wrapper script) based on the information from wrapper description file to perform input value substitution in the specified input files and for executable invocation, when executed.
- ➢ Executes a pre-defined shell script (DRM Script) in the remote machine which in turns submits the wrapper script as job to underlying resource manager of the remote resource, and waits for the output file.
- ➢ Uses Simple API for Grid Applications (SAGA) standard for transferring files between client and remote resource and launching a job in remote resource from the client machine

Some technological realizations

Other on-going activities

- **Speed the resolution**: GPGPU, FPGA, CELL
- **Abstraction of the HPC infrastructure**: OpenWP, MPI inter clusters

Outline

Summary

Main messages

- First part of the talk: Uncertainty Management is a real demand !
- Second part of the talk: HPC is mandatory to reach the objectives set by Uncertainty Management !

EADS

Summary

Naive metrics to think about Uncertainty Management

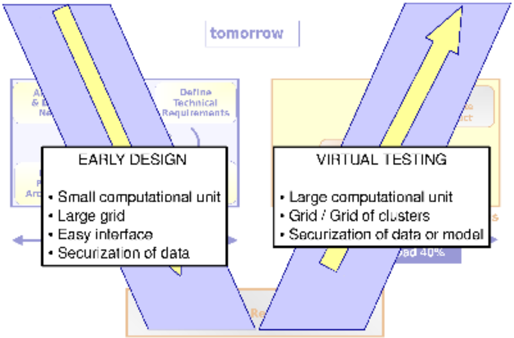
- **Engineering leverages**
  - **Choice of simulation model**

$$h \in \{h_1, \ldots, h_D\}, \text{with } h_i \in \mathcal{F}(\mathcal{X}_i \times \Theta_i, \mathcal{Y}), \ \forall i \in [1, D]$$

  - **Computational budget N**
  - **Choice of statistical laws $\mathbb{P}^X$**
- **Metrics**

$$\begin{aligned}\Delta \leq \quad & \underbrace{\mathcal{N}_{\mathcal{S}}(h^*, \tilde{h})}_{\textit{Scientific Validation}} \\ & + \underbrace{\mathcal{N}_{\mathcal{N}}(\tilde{h}, \hat{h})}_{\textit{Numerical Validation}} + \underbrace{\mathcal{N}_{\mathcal{I}}(\hat{h}, h)}_{\textit{Hardware/Software Validation}} \\ & + \underbrace{\mathcal{N}_{\mathcal{Q}}(\mathbb{P}^X_*, \mathbb{P}^X)}_{\textit{Statistical Validation}} + \underbrace{\mathcal{N}_{\mathcal{P}}(\rho(Y), \hat{\rho}_N(Y))}_{\textit{Propagation Validation}}\end{aligned}$$

Summary


tomorrow
Define Technical Requirements
EARLY DESIGN
• Small computational unit
• Large grid
• Easy interface
• Securization of data
VIRTUAL TESTING
• Large computational unit
• Grid / Grid of clusters
• Securization of data or model

Scientific challenges in HPC for UM

Numerical problems

- **Uncertainty propagation** more or less intrusive: new codes to build with HPC capabilities
- **Model reduction**: best HPC strategy to build a surrogate model
- **Inverse method**: extensive use of initial model for a stochastic optimization problem
- **Domain decomposition under uncertainty**: propagate uncertainty and couple reduced model or reduce and propagate uncertainty on coupled model ?
- **To extend the panoply of model**: To refine the numerical resolution of a problem

Scientific challenges in HPC for UM

HPC needs

- **HPC outside my resolution process**
  - Monitoring of the simulation on a heterogeneous grid
  - Best use of the HPC resources by knowing the performances of one run of the simulation model: what is the memory required? What is the best number of processors to run my computation of simulation model?
  - Fault tolerant architectures
- **HPC inside my resolution process**
  - Parallelization of some numerical steps inside the supervisor (Open TURNS like)
  - Development and use of new aleatory objects: random processes, random fields, ...
  - Open source algebra library
  - A priori/a posteriori estimation of errors
  - Fault tolerant architectures
  - Huge database manipulation

EADS

Other CHALLENGES

CULTURAL challenges

Engineers ARE NOT USED to express the uncertainty in their domain. By the way, only a few of them are trained on the subject !

- Problem to build the probabilistic criteria
- Quantification of the sources of uncertainty

A strong effort is required in basic training and professional training.

TECHNOLOGICAL challenges

The simulation tools are not adapted to evolve towards this revolution !

- Development of new algorithms
- Automatization of the computational workflow
- Is the computational budget compatible with the probabilistic criterion? Development of high performance computations capabilities.

CERTIFICATION challenges

The uncertainty management process has to be compatible with certification issues (legal responsability, safety issues, ...)

Advertisement

SAMO 2013 in Nice (July 1-4): Sensitivity analysis and design of computer experiments, http://www.gdr-mascotnum.fr/2013

CEMRACS 2013 in July-August 2013: Modelling and simulation of complex systems: stochastic and deterministic approaches. Organised by T. Lelievre, A. Nouy, N. Champagnat.

Complexity and uncertainty

> *"**Complexity** lies within the entanglement that does not allow to tackle things separately, it severs what binds groups, and produces a crippled knowledge. The problem of complexity further appears since we are part of a world, which is ruled not only by determination, stability, repetitions, or cycles, but also by outbursts and renewal. Throughout complexity, there are **uncertainties**, either empirical or theoretical, but, most of the time, both."*
>
> **Edgar Morin, Philosopher.**

EADS

Outline

- A. Antoniadis, Analyse de sensibilite, Seminaire de Toulouse 2006, www.gdr-mascotnum.fr
- Apostolakis G & Wu JS. (1993), Reliability and decision-making, Chapman & Hall London.
- Aven T. (2003), Foundations of Risk Analysis, John Wiley & Sons, Ltd.1
- Bedford T. & Cooke R. (2001), Probabilistic Risk Analysis - Foundations and Methods, Cambridge University Press.2
- R. Cont, R. Deguest, G. Scandolo (2010), Robustness and sensitivity analysis of risk measurement procedures, Quantitative Finance, Vol 10, No 6 - p593-606.
- Roger G. Ghanem and Pol D. Spanos (1991), Stochastic finite elements: A Spectral Approach, publisher Springer-Verlag - ISBN 0-387-97456-3 or 0-540-97456-3

- De Neufville R. (2003), Architecting/Designing engineering Systems using Real Options, MIT report ESD-WP-2003-01.09 (http://esd.mit.edu).
- De Neufville R. (2004), Uncertainty management for engineering systems planning and design, MIT Engineering
- EsREDA group (2008), Uncertainty in industrial practice: a guide to quantitative uncertainty management, Wiley
- Ghanem R, Spanos PD. (1991), Stochastic Finite Elements: A Spectral Approach. Springer: New York.
- Granger Morgan M. & Henrion M. (1990), Uncertainty - A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis, Cambridge University Press.3
- T. Hastie, R. Tibshirani, J. Friedman (2008) - The elements of statistical learning - Springer Series in Statistics.

- Kleijnen JPC (2009), Design and Analysis of Simulation Experiments
- Knight F.H. (1921), Risk, Uncertainty and Profit, Hart, Schaffner & Marx. Systems Monograph (http://esd.mit.edu/symposium/pdfs/monograph/uncertainty.pdf).
- P. Ladeveze (1999). Non linear Computational Structural Mechanics â New Approaches and Non-Incremental Methods of Calculation, Springer-Verlag, New York.
- Le Maitre O.P. and Knio O.M. (2010), Spectral Methods for Uncertainty Quantification With Applications to Computational Fluid Dynamics. Scientific Computation.
- Lucor D. (2006), Presentation at CEMRACS'06
- Maday Y., Reduced Basis Method for the Rapid and Reliable Solution of Partial Differential Equations, hal-00112152, version 1.
- GDR CNRS MASCOT-NUM, http://www.gdr-mascotnum.fr

EADS

- P. Massart (2007), Concentration inequalities and model selection: Ecole d'Ete de Probabilites de Saint-Flour XXXIII-2003. Springer Verlag.
- H.G. Matthies (2007), Uncertainty Quantification with Stochastic Finite Elements. Encyclopedia of Computational Mechanics.
- Managing Uncertainty in Complex Models (MUCM), http://www.mucm.ac.uk/
- Open Source initiative in uncertainty managempent, http://www.openturns.org
- A. Owen, J.R. Koehler, (1996), Computer Experiments
- High Performance Computing & Uncertainty Quantification, CEA-EDF-INRIA summer school 2011

- N. Rachdi, J-C Fort, T. Klein (2010), Risk bounds for new M-estimation problems, http://hal.archives-ouvertes.fr/hal-00537236.
- N. Rachdi (2011), PhD thesis from University Paul Sabatier of Toulouse.
- J Sacks et al (1989), Design and Analysis of Computer Experiments, Technometrics
- Saltelli A., K. Chan, and E.M. Scott (2000). Sensitivity Analysis. Chichester: Wiley
- Santner, Thomas (2003), The design and analysis of computere experiments -ISBN 0-387-95420-1
- Sobol, I. M. (1993). Sensitivity estimates for non linear mathematical models. Mathematical Modeling and Computational Experiment 1, 407-414

de Weck O. & Eckert C. (2007), A classification of uncertainty for early product and system design, MIT report ESD-WP-2007-10 (http://esd.mit.edu/WPS).