



Metamodels and polynomial chaos expansions

Géraud Blatman, EDF R&D

HPC and Uncertainty Treatment
Examples with Open TURNS and URANIE

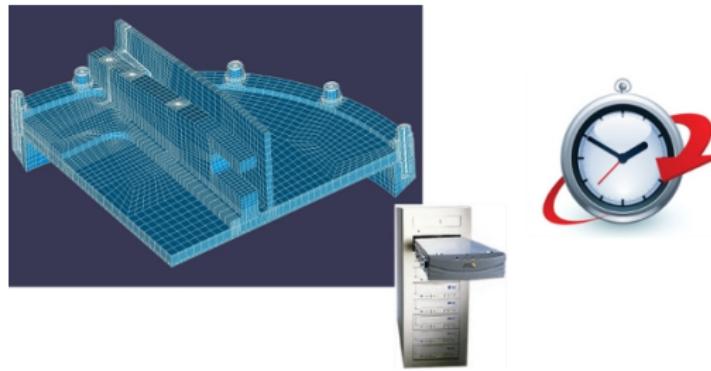
Airbus Group – CEA – EDF – IMACS – Phimeca

PRACE Advanced Training Center
May 19th 2016



Inaffordable engineering analyses for real world problems

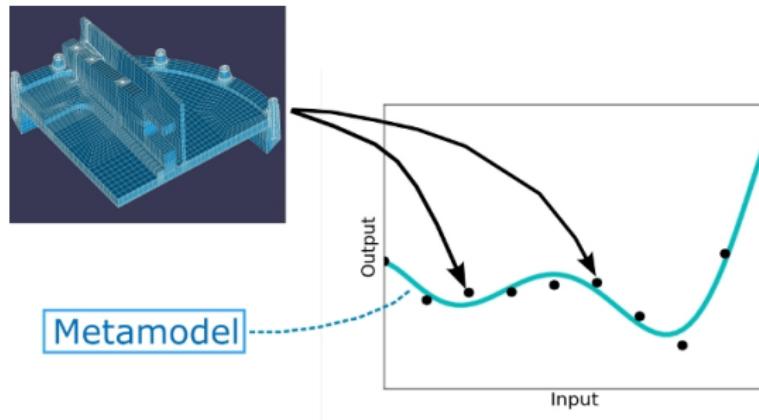
For complex models, a single simulation may take several hours or days to complete.



This makes tasks as optimization, design space exploration, uncertainty propagation and sensitivity analysis impossible in practice.

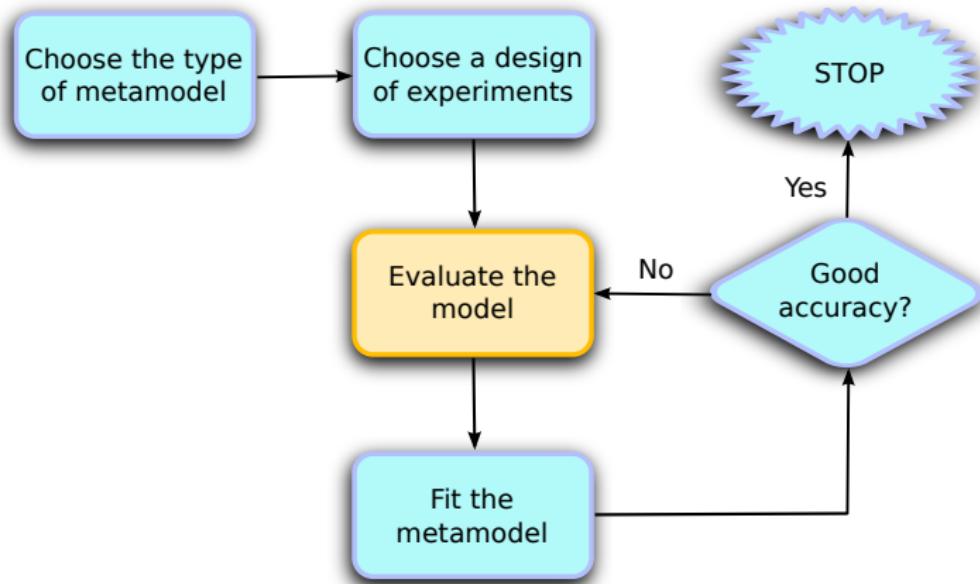
A remedy named metamodel

Construct approximation models known as **metamodels** that mimic the behaviour of the simulator while being cheaper to evaluate.



Then use the surrogate instead of the simulator in order to perform computationally expensive analyses.

The major steps for constructing a metamodel



Outline

- 1 Polynomial chaos
- 2 A short introduction to Gaussian process modelling
- 3 Conclusions

Outline

1 Polynomial chaos

- Unidimensional case with a uniform variable
- Unidimensional case with a non uniform variable
- Multidimensional case
- Validation and stepwise construction of an optimal chaos approximation

2 A short introduction to Gaussian process modelling

3 Conclusions

Outline

1 Polynomial chaos

- Unidimensional case with a uniform variable
- Unidimensional case with a non uniform variable
- Multidimensional case
- Validation and stepwise construction of an optimal chaos approximation

2 A short introduction to Gaussian process modelling

3 Conclusions

Instructive example

Model: $Y = f(X) = X \sin X$, $X \sim \mathcal{U}([0, 10])$

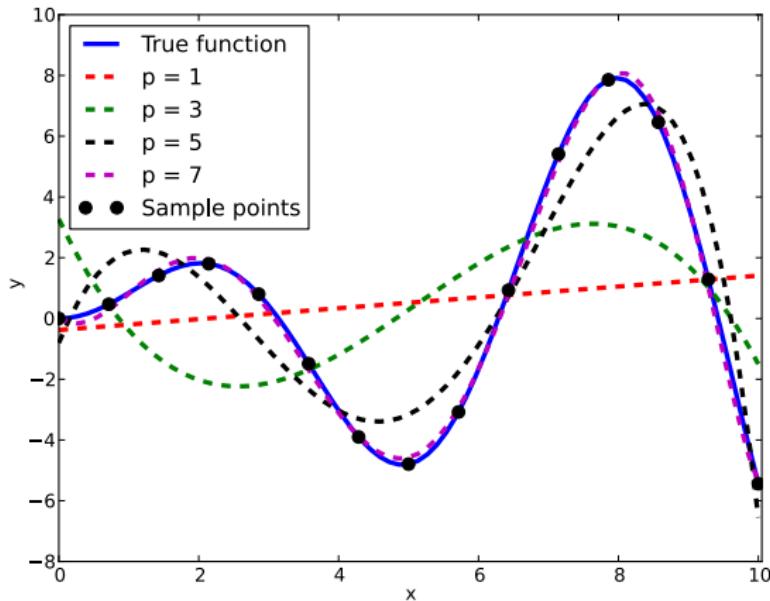
Polynomial metamodel: $\tilde{f}_{\boldsymbol{a}}(X) = \sum_{j=0}^p a_j X^j$

Coefficients estimation by least squares:

- Design of experiments: n points $x^{(i)}$ evenly distributed in $[0, 10]$ → Vecteurs $\mathcal{X} = (x^{(1)}, \dots, x^{(n)})$ et $\mathcal{Y} = (y^{(1)}, \dots, y^{(n)})$
- Design matrix: $\boldsymbol{\pi} = (x^{(i)^j}, i = 1, \dots, n, j = 0, \dots, p)$
- Coefficients estimation:
$$\hat{\boldsymbol{a}} = (\boldsymbol{\pi}^\top \boldsymbol{\pi})^{-1} \boldsymbol{\pi}^\top \mathcal{Y}$$

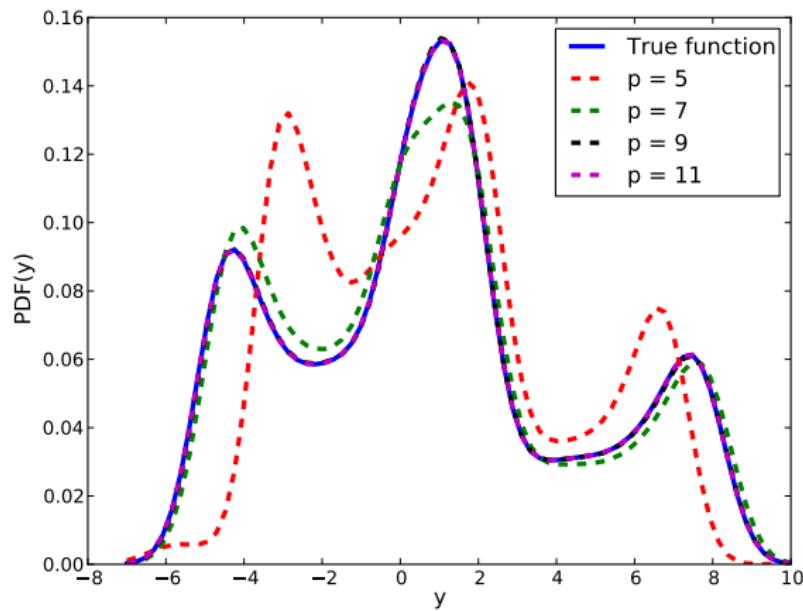
Convergence with respect to the degree p

Design size: $n = 15$



Convergence in terms of probability density function (PDF)

Evaluation of the PDF of each metamodel $\tilde{f}_a(X)$ based on 50,000 realizations of X



Estimation of the moments of Y

Naive strategy: Evaluate $\tilde{f}_a(X)$ for a large number of realizations of X and compute the sample moments of the set of results

Drawback: The results are affected by a sampling error

Alternative: Derive analytically the moments by using centered and orthonormal polynomials

Centered and orthonormal polynomials

- Let $(\pi_j)_{j=0,1,2,\dots}$ be a family of polynomials
- The polynomials are said to be centered w.r.t the random variable ξ IFF :

$$\mathbb{E} [\pi_i(\xi)] \equiv \int_{\mathcal{D}_\xi} \pi_i(u) f_\xi(u) du = 0 \quad \forall i \geq 1$$

where \mathcal{D}_ξ and $f_\xi(u)$ are the support the PDF of ξ

- They are said to be orthonormal w.r.t the random variable ξ IFF :

$$\mathbb{E} [\pi_i(\xi) \pi_j(\xi)] \equiv \int_{\mathcal{D}_\xi} \pi_i(u) \pi_j(u) f_\xi(u) du = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

Case of a uniform variable – Legendre polynomials

- Let the random variable : $\xi \sim \mathcal{U}([-1, 1])$
- The orthonormal polynomials w.r.t ξ are the Legendre polynomials:

$$\pi_0(x) = 1 , \quad \pi_1(x) = \sqrt{3}x , \quad \pi_2(x) = \frac{\sqrt{5}}{2}(3x^2 - 1) , \quad \dots$$

$$\mathbb{E}[\pi_i(\xi)] = 0 , \quad \forall i \geq 1 \quad \text{et} \quad \mathbb{E}[\pi_i(\xi) \pi_j(\xi)] = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

Instructive example – Legendre approximation

- Model: $Y = f(X) = X \sin X$, $X \sim \mathcal{U}([0, 10])$
- Change of variable: $\xi \equiv \frac{X}{5} - 1$, $\xi \sim \mathcal{U}([-1, 1])$
- Metamodel called “polynomial chaos”:

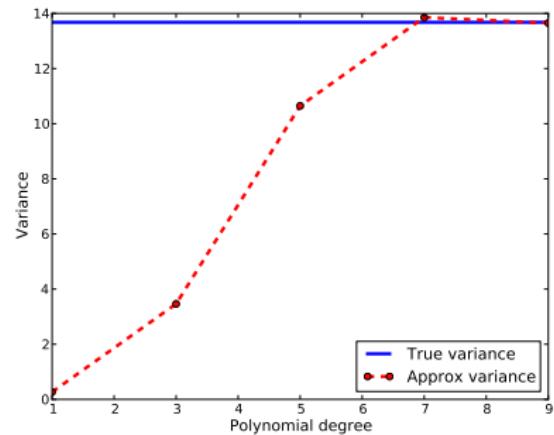
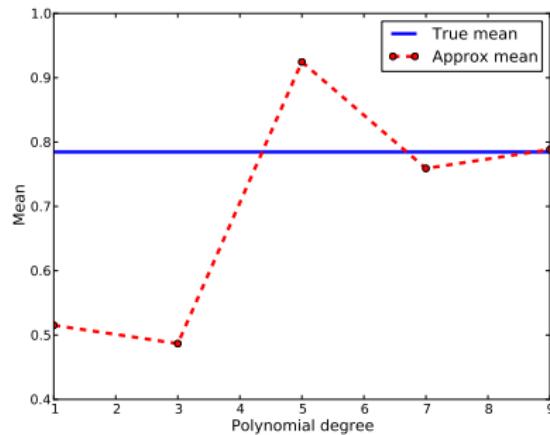
$$\tilde{f}_{\mathbf{a}}(X) = \tilde{h}_{\mathbf{a}}(\xi) = \sum_{j=0}^p a_j \pi_j(\xi)$$

where the π_j 's are the Legendre polynomials and the a_j 's are the least squares coefficients

- Closed-form formulas of the moments:

$$\mathbb{E} [\tilde{f}_{\mathbf{a}}(X)] = a_0, \quad \text{Var} [\tilde{f}_{\mathbf{a}}(X)] = \sum_{j=1}^p a_j^2$$

Convergence of the moments



Accurate moment estimates from $p = 7$

Outline

1 Polynomial chaos

- Unidimensional case with a uniform variable
- **Unidimensional case with a non uniform variable**
- Multidimensional case
- Validation and stepwise construction of an optimal chaos approximation

2 A short introduction to Gaussian process modelling

3 Conclusions

Instructive example with a Gaussian variable

Model: $Y = f(X) = X \sin X$, $X \sim \mathcal{N}(5, 1)$

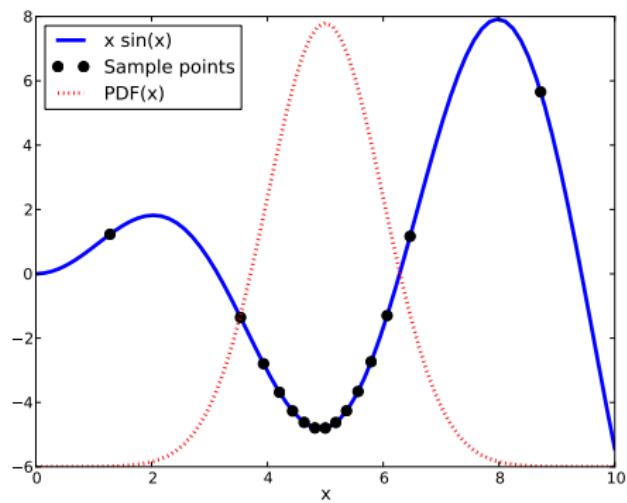
Change of variable : $\xi \equiv X - 5$, $\xi \sim \mathcal{N}(0, 1)$

Polynomial chaos: $\tilde{f}_{\mathbf{a}}(X) = \tilde{h}_{\mathbf{a}}(\xi) = \sum_{j=0}^p a_j \pi_j(\xi)$

where the π_j 's are the **Hermite polynomials** which are centered and orthonormal w.r.t ξ

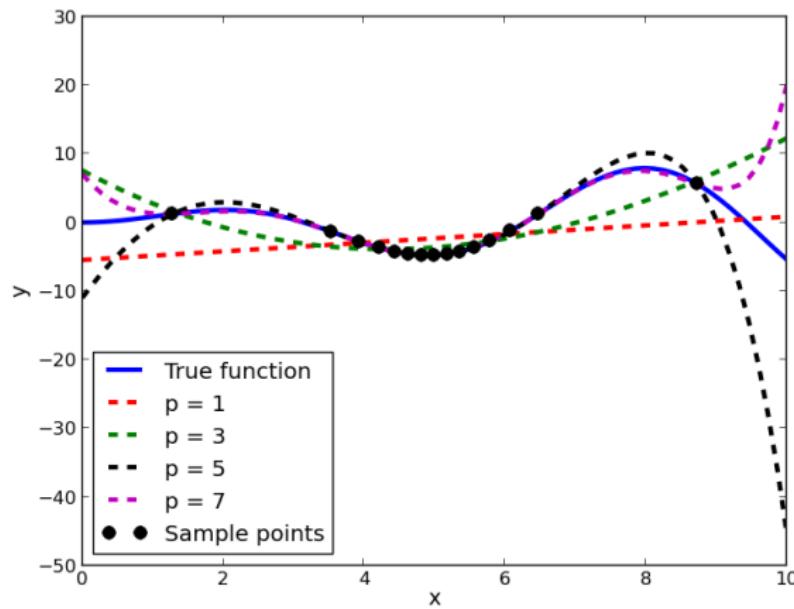
Design of experiments to compute the coefficients

- Sampling of the $x^{(i)}$'s according to the PDF $\mathcal{N}(5, 1)$

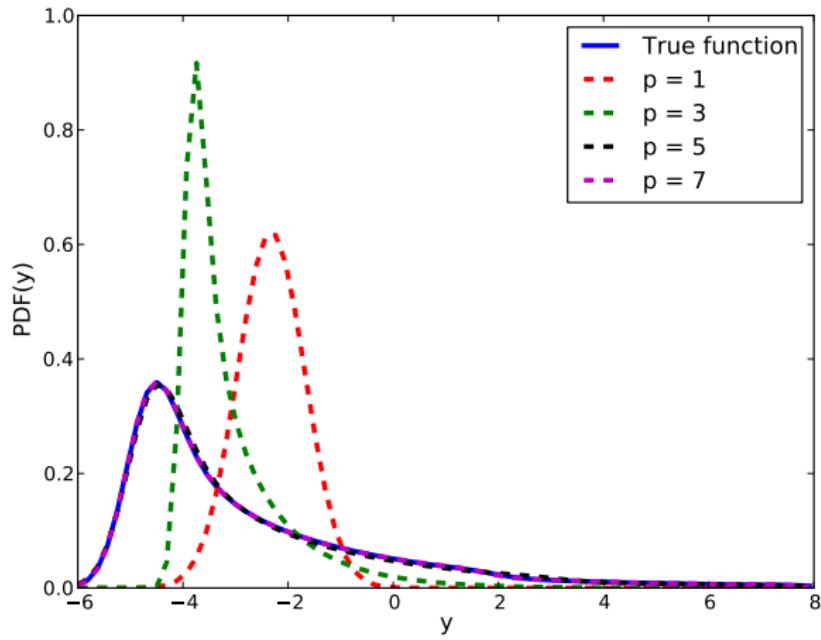


- Change of variable for each realization: $\xi^{(i)} = x^{(i)} - 5$

Approximation of the output



Approximation of the output PDF



Correspondences usual PDFs - orthonormal polynomials

PDF of ξ	Support	Polynomials
Normal (Gaussian)	\mathbb{R}	Hermite
Uniform	$[-1, 1]$	Legendre
Gamma	$(0, \infty)$	Laguerre
Chebyshev	$(-1, 1)$	Chebyshev
Beta	$(-1, 1)$	Jacobi

Change of variable

General formula

$$\xi = F_{\xi}^{-1}[F_X(X)]$$

where F_{ξ} and F_X are the (known) cumulative density functions of ξ and X

Guidelines for choosing the couple (ξ , polynomials)

- If the type of the X PDF belongs to the table, select the ξ with the same PDF (e.g. $\mathcal{N}(210, 18) \rightarrow \mathcal{N}(0, 1)$)
- Otherwise, select a ξ with the same type of support (infinite, semi-bounded or bounded) than X

Outline

1 Polynomial chaos

- Unidimensional case with a uniform variable
- Unidimensional case with a non uniform variable
- **Multidimensional case**
- Validation and stepwise construction of an optimal chaos approximation

2 A short introduction to Gaussian process modelling

3 Conclusions

Instructive example

Model:
$$Y = f(X_1, X_2) = X_1 \sin X_2$$
$$X_1 \sim \mathcal{U}([0, 10]) \quad , \quad X_2 \sim \mathcal{LN}(5, 1)$$

Changes of variable :

- $\xi_1 = X_1/5 - 1$, $\xi_1 \sim \mathcal{U}([-1, 1]) \rightarrow$ Legendre polynomials
- $\xi_2 = \exp(X_2) - 5$, $X_2 \sim \mathcal{N}(0, 1) \rightarrow$ Hermite polynomials

Construction of a bivariate orthonormal basis

- Legendre polynomials:

$$\pi_0^{(1)}(\xi_1) = 1 \quad , \quad \pi_1^{(1)}(\xi_1) = \sqrt{3}\xi_1 \quad , \quad \pi_2^{(1)}(x) = \frac{\sqrt{5}}{2}(3\xi_1^2 - 1) \quad , \quad \dots$$

- Hermite polynomials:

$$\pi_0^{(2)}(\xi_2) = 1 \quad , \quad \pi_1^{(2)}(\xi_1) = \xi_2 \quad , \quad \pi_2^{(2)}(\xi_2) = \frac{\sqrt{2}}{2}(\xi_2^2 - 1) \quad , \quad \dots$$

- Construction of bivariate polynomials:

$$\psi_{0,0}(\xi_1, \xi_2) = \pi_0^{(1)}(\xi_1) \times \pi_0^{(2)}(\xi_2) = 1$$

$$\psi_{1,0}(\xi_1, \xi_2) = \pi_1^{(1)}(\xi_1) \times \pi_0^{(2)}(\xi_2) = \sqrt{3}\xi_1$$

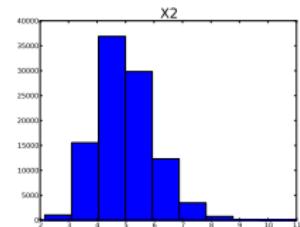
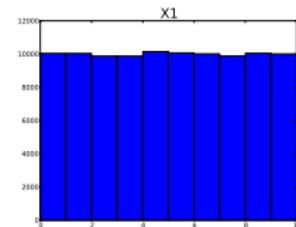
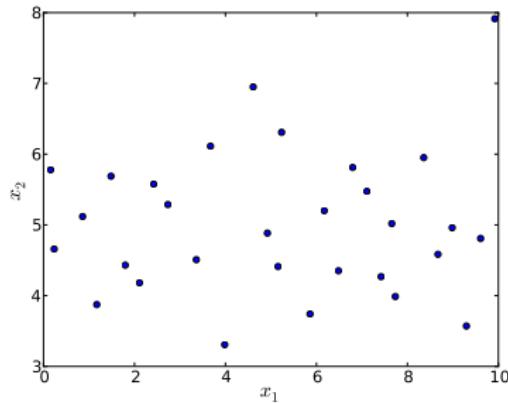
$$\psi_{1,2}(\xi_1, \xi_2) = \pi_1^{(1)}(\xi_1) \times \pi_2^{(2)}(\xi_2) = \frac{\sqrt{6}}{2}\xi_1(\xi_2^2 - 1)$$

Fitting of a polynomial chaos approximation

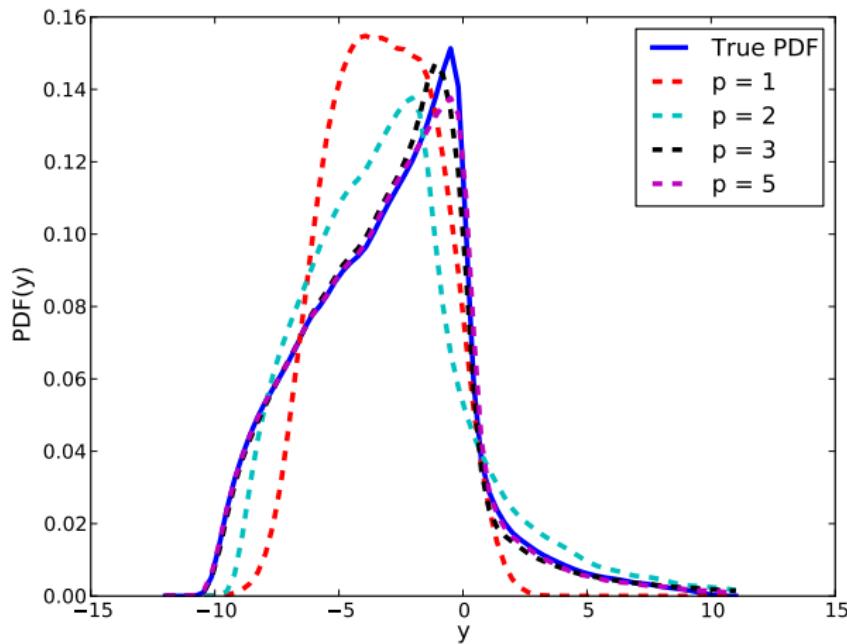
“Hybrid” Legendre-Hermite polynomial chaos:

$$\tilde{f}_{\boldsymbol{a}}(X_1, X_2) = \tilde{h}_{\boldsymbol{a}}(\xi_1, \xi_2) = \sum_{i_1+i_2 \leq p} a_{i_1, i_2} \psi_{i_1, i_2}(\xi_1, \xi_2)$$

Quasi-random design of experiments ($n = 30$):



Convergence of the output PDF



Multivariate polynomial chaos

- Consider the model:

$$Y = f(\mathbf{X}) \quad , \quad \dim(\mathbf{X}) = M \quad , \quad \dim(Y) = 1$$

- Assumption: independence of the components of \mathbf{X}
- Approximation by tensorizing the 1D scheme:

$$\tilde{f}_{\mathbf{a}}(\mathbf{X}) = \sum_{i_1 + \dots + i_M \leq p} a_{i_1, \dots, i_M} \pi_{i_1}^{(1)}(X_1) \times \dots \times \pi_{i_M}^{(M)}(X_M)$$

The centering and orthonormality properties still hold
→ Closed-form formulas of the moments
and the Sobol sensitivity indices

Derivation of the Sobol sensitivity indices

Sobol sensitivity indices: Quantify the part of the variance of Y that is due to the variance of each variable X_i

Definition:

$$S[X_i] = \frac{\text{Var}[\mathbb{E}[Y|X_i]]}{\text{Var}[Y]}$$

Direct estimation from model evaluations can reveal inaffordable!

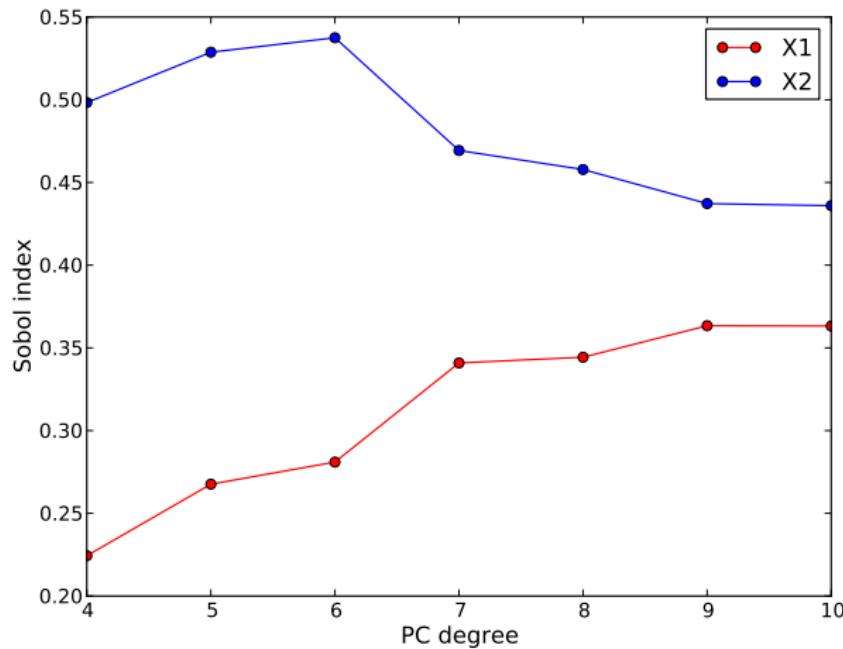
Estimation of the numerator from the chaos proxy :

$$\boxed{\text{Var}[\mathbb{E}[Y|X_i]] \approx \sum_{j \in \mathcal{I}_i} a_j^2}$$

where \mathcal{I}_i is set of the “multi-indices” j such that $j_k = 0$ if $k \neq i$

Sensitivities to interactions can also be derivated!

Instructive example – Convergence of the Sobol indices



Outline

1 Polynomial chaos

- Unidimensional case with a uniform variable
- Unidimensional case with a non uniform variable
- Multidimensional case
- Validation and stepwise construction of an optimal chaos approximation

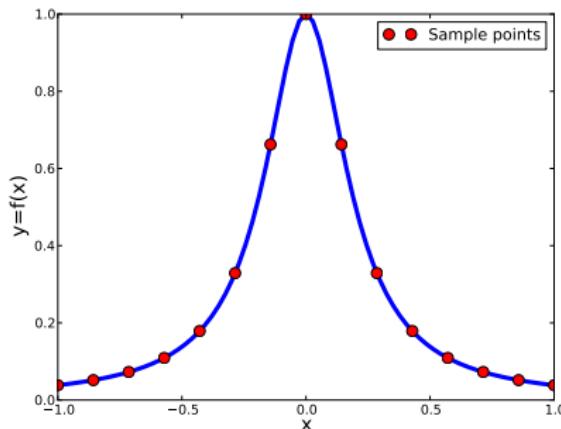
2 A short introduction to Gaussian process modelling

3 Conclusions

Runge function

Model: $Y = f(X) = \frac{1}{1 + 25X^2}$, $X \equiv \xi \sim \mathcal{U}([-1, 1])$

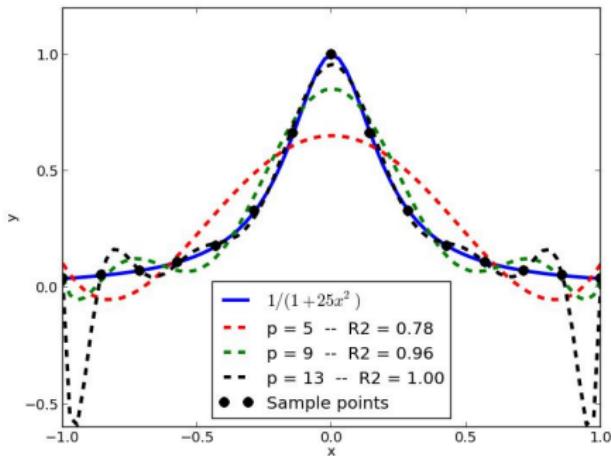
Design of experiments: $n = 15$ points $x^{(i)}$ evenly located over $[-1, 1]$



Legendre chaos approximation : $Y \approx \tilde{f}_a(\xi) = \sum_{j=0}^p a_j \pi_j(\xi)$

The R^2 accuracy estimator

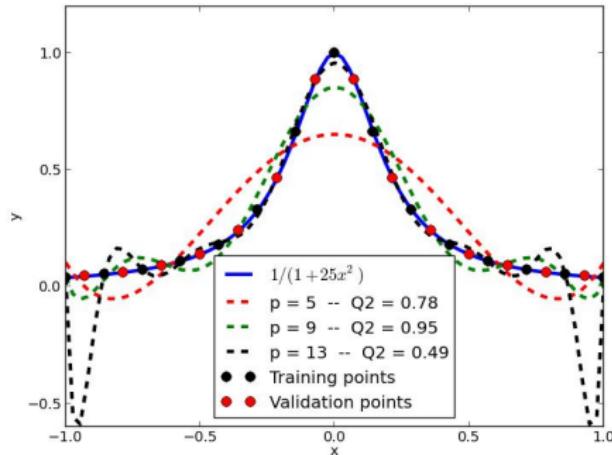
$$R^2 = 1 - Err \quad , \quad Err \propto \sum_{i=1}^n \left(f(\xi^{(i)}) - \tilde{f}_a(\xi^{(i)}) \right)^2$$



Problem : R^2 increases with the error (overfitting) \rightarrow Not reliable estimator

Alternative: use of a validation point set

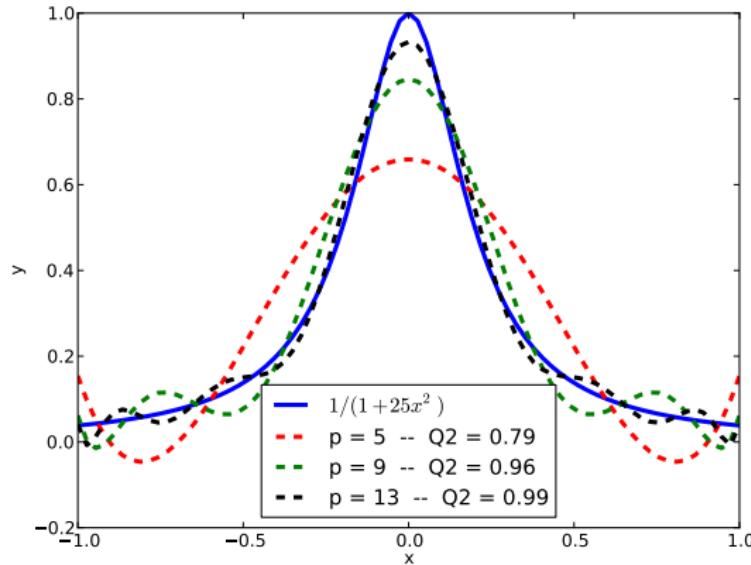
- Sampling of m new points $\{x'^{(1)}, \dots, x'^{(m)}\}$
- Calculation of the R^2 counterpart from these points → Coefficient Q^2



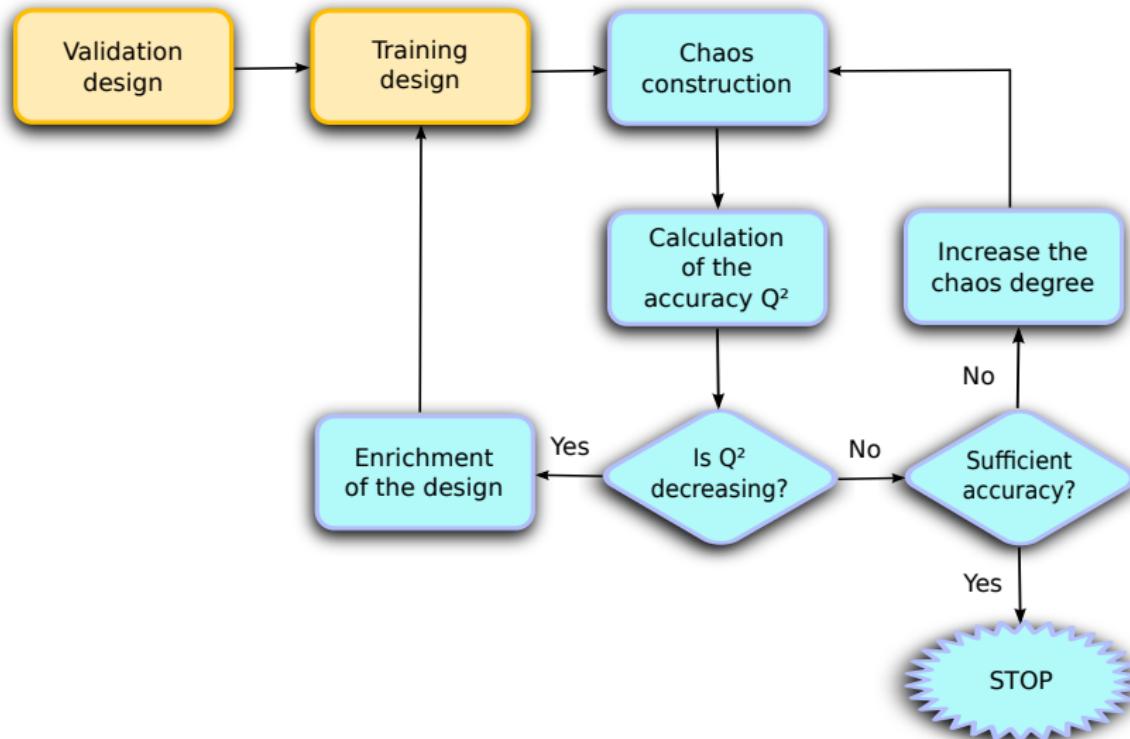
Q^2 “detects” overfitting → More reliable estimator

Improving the approximation by adding design points

Addition of 15 new points in the design of experiments



Adaptive construction of an “optimal” polynomial chaos



Some remarks

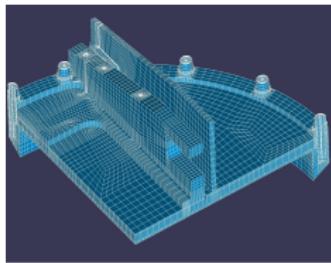
- There exists other methods to estimate the coefficients ([numerical integration](#))
- Ill-conditioned least squares problem when the design size n is close to the number of chaos coefficients ([overfitting](#))
 - Obtaining an accurate approximation in high dimensions (≥ 10 input parameters) is uneasy

Solution: Use of [sparse least squares](#) ([Least Angle Regression](#))

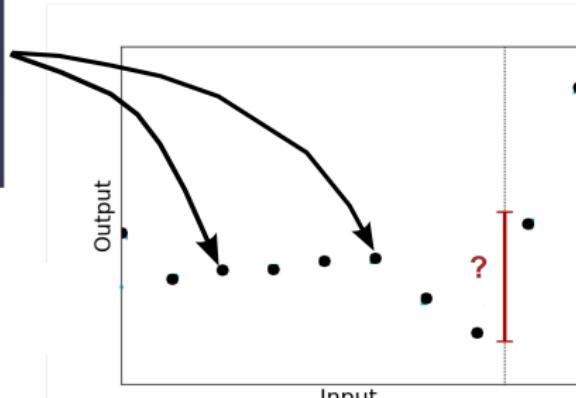
Outline

- 1 Polynomial chaos
- 2 A short introduction to Gaussian process modelling
- 3 Conclusions

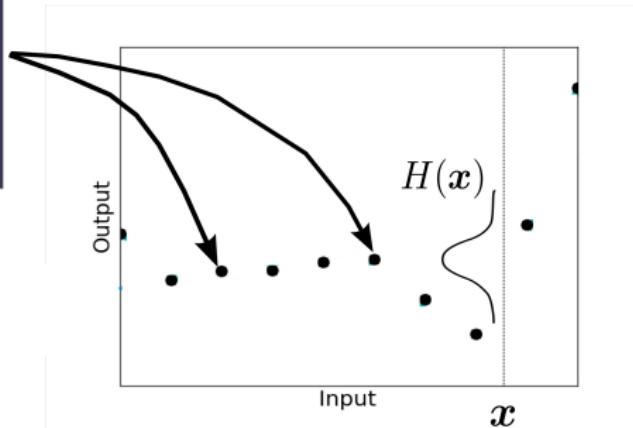
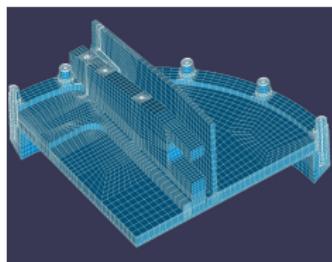
Prediction at a new point (1)



Output value at a new location?



Prediction at a new point (2)



Assumption: The response is a realization of a Gaussian random variable whose moments depend on the design points

Gaussian process assumption

The model output is a realization of a Gaussian random process of the form :

$$H(\boldsymbol{x}, \omega) = \boxed{\boldsymbol{r}(\boldsymbol{x}) \cdot \boldsymbol{\beta}} + \boxed{Z(\boldsymbol{x}, \omega)}$$

Trend (deterministic)

Linear regression
on a fixed basis

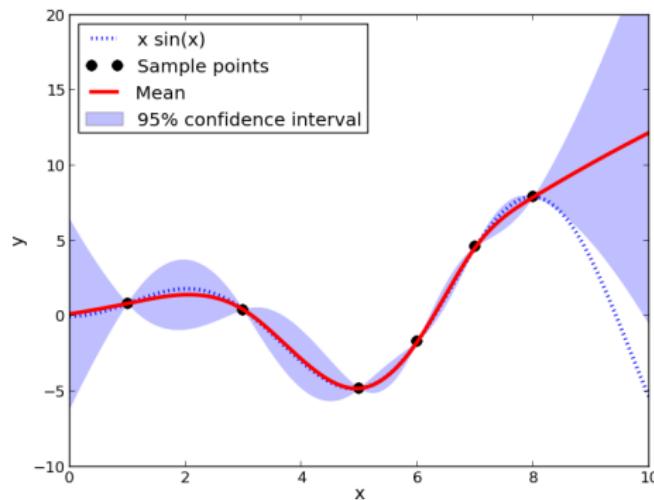
Random fluctuations

Zero-mean and
stationary
Gaussian process

$$\text{Cov}_Z(\boldsymbol{x}, \boldsymbol{x}') = \sigma^2 \rho(\|\boldsymbol{x} - \boldsymbol{x}'\|)$$

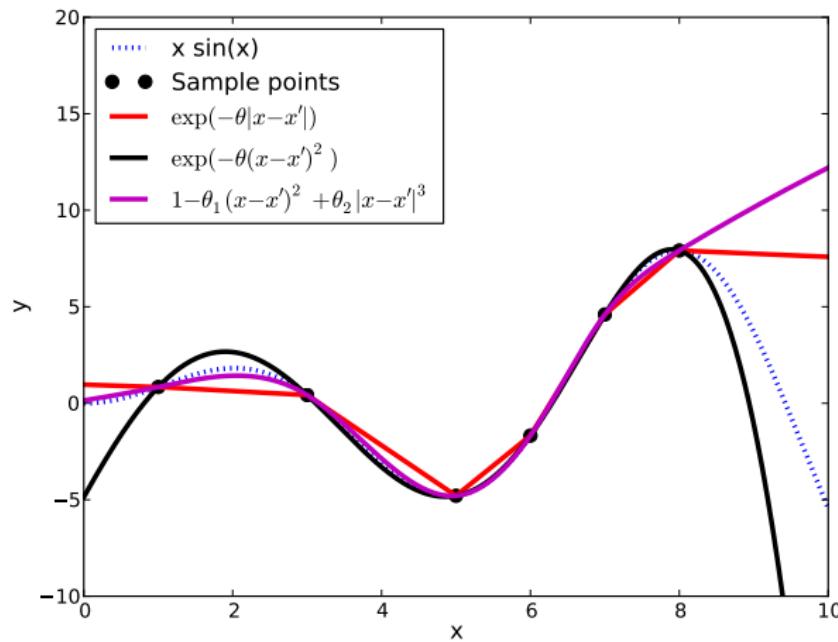
Instructive example — $y = f(x) = x \sin(x)$

$$H(x, \omega) = \beta_0 + \beta_1 x + Z(x, \omega), \quad \text{Cov}_Z(x, x') = \sigma^2 e^{-\theta(x-x')^2}$$



- The mean is used as a **metamodel** (interpolator)
- The prediction bands are used as **error indicators**

Influence of the autocorrelation function



The autocorrelation function type has an effect on the metamodel

smoothness

Expressions of the mean and the variance

Notations: $\mathbf{k}(\mathbf{x}) \equiv \{\rho(\mathbf{x}, \mathbf{x}^{(1)}), \dots, \rho(\mathbf{x}, \mathbf{x}^{(N)})\}^\top$

$\mathbf{R} \equiv (r_j(\mathbf{x}^{(i)}))_{1 \leq i, j \leq N}$, $\mathbf{K} \equiv (\rho(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}))_{1 \leq i, j \leq N}$

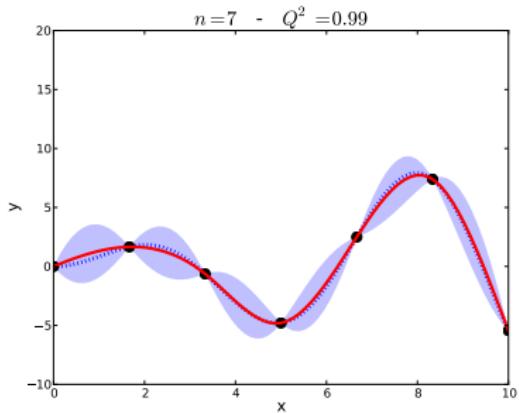
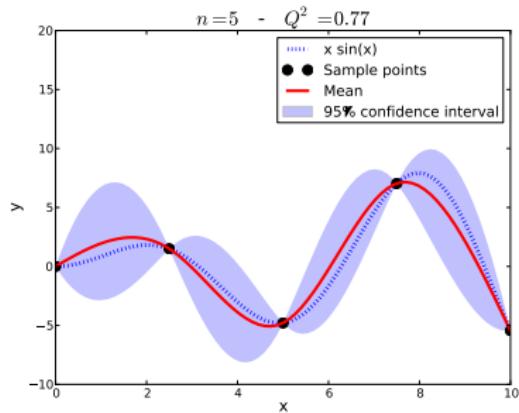
Mean:
$$\mu(\mathbf{x}) = \mathbf{r}^\top(\mathbf{x})\boldsymbol{\beta} + \mathbf{k}^\top(\mathbf{x})\mathbf{K}^{-1}(\mathcal{Y} - \mathbf{R}\boldsymbol{\beta})$$

Variance:
$$\sigma^2(\mathbf{x}) = \sigma^2 - \mathbf{k}^\top(\mathbf{x})\mathbf{K}^{-1}\mathbf{k}^\top(\mathbf{x})$$

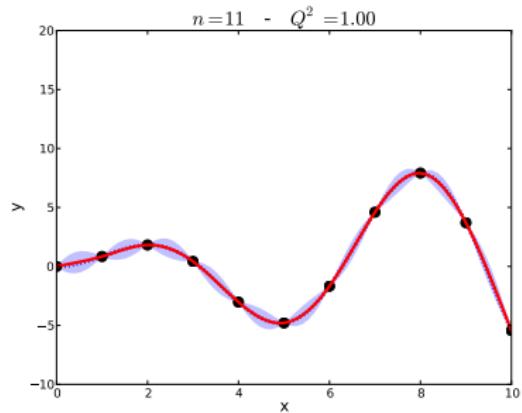
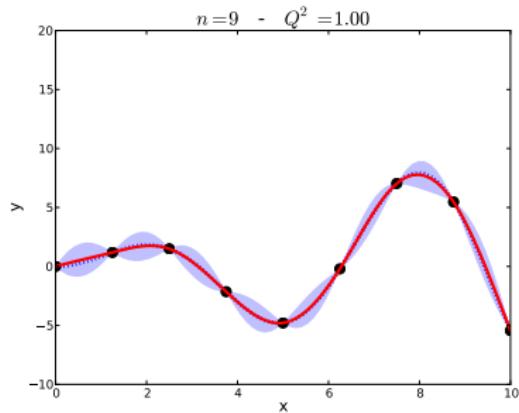
Parameter fitting

- To apply the previous formulas, the parameters (β, σ, θ) have to be estimated from the design points
- Estimation by the maximum likelihood method

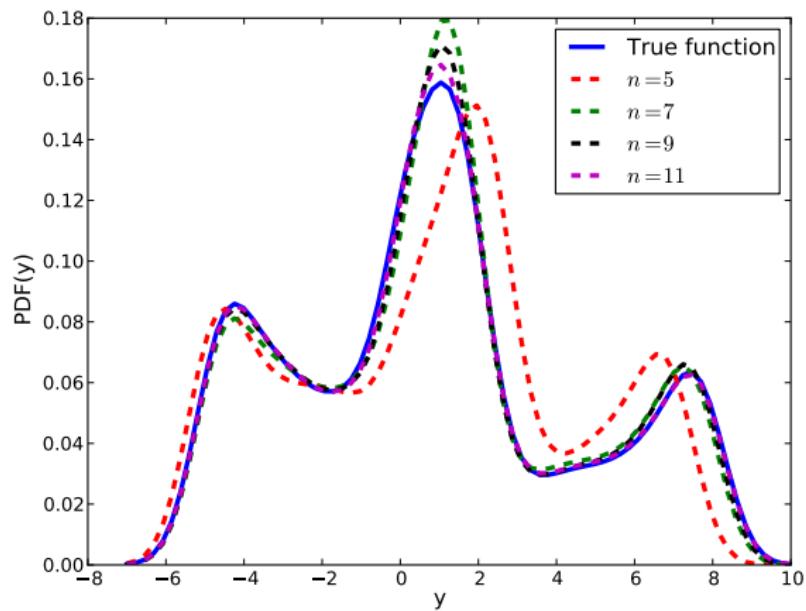
Approximation of $f(x) = x \sin(x)$ – Convergence (1)



Approximation of $f(x) = x \sin(x)$ – Convergence (2)



Approximation of $f(x) = x \sin(x)$ – Convergence (3)



Outline

- 1 Polynomial chaos
- 2 A short introduction to Gaussian process modelling
- 3 Conclusions

Polynomial chaos or Gaussian process?

	Polynomial chaos	Gaussian process
Pros	Closed-form formulas of the moments and the sensitivities Algorithmic simplicity	Error indicators (prediction intervals) Possibility of refining optimally the design of experiments
Cons	No “local” error indicator	Uneasy parameter estimation for dimensions ≥ 15

Some computer tools (non exhaustive list!)

Polynomial chaos

Gaussian process

OpenTURNS (Python)

UQLab (Matlab)

URANIE

chaospy (Python)

Scikits (Python)

UQToolbox (Python)

DACE (Matlab)

NISP (Scilab)

STK (Matlab)

DiceKriging (R)

A few messages...

- Polynomial chaos applied to several industrial studies (civil engineering, non destructive testing, etc.)
- For a given problem, ideally test several types of metamodels
- Check that the designs are consistent with the input parameter PDFs!
- Assess the metamodel accuracy prior to computing quantities of interest!
- Metamodelling may not work for highly nonlinear models

Thank you for your attention!