Functional data reduction issue for robust inversion

Reda El Amri [1 2], Céline Helbert [3], O. Lepreux [2], Miguel Munoz Zuniga [2],
Clémentine Prieur [1], Delphine Sinoquet [2]

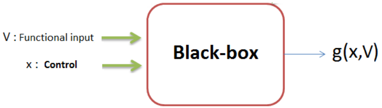
Workshop on Design of Experiments
CIRM (Marseilles, France) April 30 - May 4, 2018

May 4, 2018

1. Grenoble Alpes University
2. French Institute of Petroleum
3. Ecole Centrale de Lyon

Table of Contents

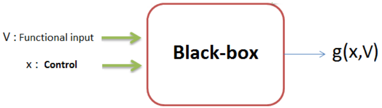
The framework : automotive pollution control



$$g : \begin{array}{ccc} D \times \mathcal{F} & \to & \mathbb{R} \\ (\mathbf{x}, \mathbf{v}) & \mapsto & g(\mathbf{x}, \mathbf{v}) \end{array}$$

where $D \subset \mathbb{R}^p$ is a compact and $\mathcal{F}$ a functional space.

The framework : automotive pollution control



$$\begin{array}{rcccc} g & : & D \times \mathcal{F} & \to & \mathbb{R} \\ & & (\mathbf{x}, \mathbf{v}) & \mapsto & g(\mathbf{x}, \mathbf{v}) \end{array}$$

where $D \subset \mathbb{R}^p$ is a compact and $\mathcal{F}$ a functional space.

Probabilistic description of uncertainty : $\mathbf{V}$ is a random variable valued in $\mathcal{F}$.

For a fixed $s \in \mathbb{R}$, define

$$\begin{array}{lcl} \Gamma^*_{a.s.} & := & \{\mathbf{x} \in D \ s.t. \ g(\mathbf{x}, \mathbf{V}) \leq s \quad \textit{almost surely}\}, \\ \Gamma^*_{\alpha} & := & \{\mathbf{x} \in D \ s.t. \ \mathbb{P}(g(\mathbf{x}, \mathbf{V}) \leq s) \ \geq \ 1 - \alpha\}, \\ \Gamma^* & := & \{\mathbf{x} \in D \ s.t. \ f(x) = \mathbb{E}[g(\mathbf{x}, \mathbf{V})] \leq s\} \quad := f^{-1}(T) \text{ , where } T = (-\infty, s]. \end{array}$$

Aim : estimate Γ^*, the excursion set of f below t.

The framework

Model : $g : D \times \mathcal{F} \to \mathbb{R}$ with $D \subset \mathbb{R}^p$ and $\mathcal{F}$ a functional space.

Objective : estimate $\Gamma^* \subset \mathbb{R}^p$ from n evaluations of function $f(\cdot) = \mathbb{E}[g(\cdot, \mathbf{V})]$, $\mathbf{f}_n = (f(\mathbf{x}^1), f(\mathbf{x}^2), ..., f(\mathbf{x}^n))$.

Issues :

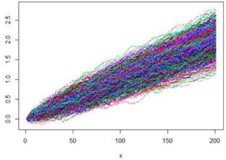
- each evaluation $f(\mathbf{x})$ requires the estimation of $\mathbb{E}[g(\mathbf{x}, \mathbf{V})]$,
- $\mathbf{V}$ is known through κ realizations $\mathbf{v}^1, \ldots \mathbf{v}^\kappa$,
- each evaluation $g(\mathbf{x}, \mathbf{v})$ is costful.

An analytical example (1/5)

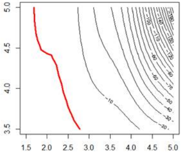
Let $g : D \times \mathcal{F} \to \mathbb{R}$ defined as

$$g(\mathbf{x}, \mathbf{V}) = |0.1 \cos(x_1 \max_t V_t) \sin(x_2).(x_1 + x_2 \min_t V_t)^2|. \int (30 + V_t)^{\frac{x_1 \cdot x_2}{20}} \mathrm{d}t. \max_t V_t$$

with $\mathbf{x} = (x_1, x_2) \in D \subset \mathbb{R}^2$ and $\mathbf{V} = (V_t, t \geq 0)$ a BM with constant drift equal to 2.



$$\begin{aligned}\Gamma^* &= \{\mathbf{x} \in [1.5, 5] \times [3.5, 5] \ : \ f(\mathbf{x}) \leq 1.2\} \\ &= \{\mathbf{x} \in [1.5, 5] \times [3.5, 5] \ : \ \mathbb{E}\left[g(\mathbf{x}, \mathbf{V})\right] \leq 1.2\}\end{aligned}$$



Bayesian Approach / Kriging [4]

f is seen as a realization of a Gaussian Process $(Z_{\mathbf{x}})_{\mathbf{x}\in D}$ with prior mean m and covariance kernel k.

Γ^* is a realization of $\Gamma := \{\mathbf{x} \in D : Z(\mathbf{x}) \leq s\}$.

Let $\mathcal{X}_n = (\mathbf{x}^1, \ldots, \mathbf{x}^n)$. Given the evaluations $\mathbf{f}_n = f(\mathcal{X}_n)$, the posterior field $Z|Z(\mathcal{X}_n) = \mathbf{f}_n$ follows a Gaussian distribution with mean and covariance kernel :

$$\begin{aligned} m_n(\mathbf{x}) &= m(\mathbf{x}) + k(\mathbf{x}, \mathcal{X}_n)k(\mathcal{X}_n, \mathcal{X}_n)^{-1}\left(\mathbf{f}_n - m(\mathcal{X}_n)\right) \\ k_n(\mathbf{x}, \mathbf{y}) &= k(\mathbf{x}, \mathbf{y}) - k(\mathbf{x}, \mathcal{X}_n)k(\mathcal{X}_n, \mathcal{X}_n)^{-1}k(\mathcal{X}_n, \mathbf{y}) \end{aligned}$$

What about the probability distribution of $\Gamma|Z(\mathcal{X}_n) = \mathbf{f}_n$?

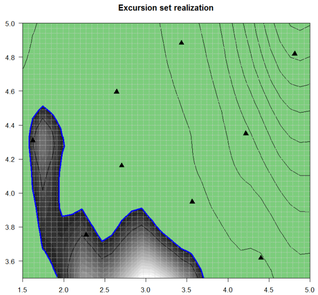
4. Clément CHEVALIER (2013). ≪ Fast uncertainty reduction strategies relying on Gaussian process models ≫. Thèse de doct. Citeseer

An analytical example (2/5)

What about the probability distribution of $\Gamma | Z(\mathcal{X}_n) = \mathbf{f}_n$?

Let us consider :

- priors on the Gaussian process : constant mean function, 5/2 Matern covariance kernel ;
- initial DoE $\mathcal{X}_n = (\mathbf{x}^1, \ldots, \mathbf{x}^n)$, $n = 9$ (black triangles).



How to summarize the probability distribution of $\Gamma | Z(\mathcal{X}_n) = \mathbf{f}_n$?

What are, for random sets, the notions of

- expectation,
- deviation.

Remark : let μ be a Borel measure on D. Then,

$$\mathbb{E}\left(\mu(\Gamma) | Z(\mathcal{X}_n) = \mathbf{f}_n\right) = \int \mathbb{P}_n(\mathbf{x} \in \boldsymbol{\Gamma}) \mu(d\mathbf{x})$$

with

$$\mathbb{P}_n(\mathbf{x} \in \Gamma) = \mathbb{P}\Big(Z(\mathbf{x}) \leq s | Z(\mathcal{X}_n) = \mathbf{f}_n\Big).$$

Theory of random sets, Vorob'ev quantiles [5] [6]

The conditional coverage function of $\boldsymbol{\Gamma}$ is defined as

$$p_n : \mathbf{x} \in D \longrightarrow \mathbb{P}_n(\mathbf{x} \in \Gamma) = \mathbb{P}\left(Z(\mathbf{x}) \leq s \mid Z(\mathcal{X}_n) = \mathbf{f}_n\right) \quad \in [0,1].$$

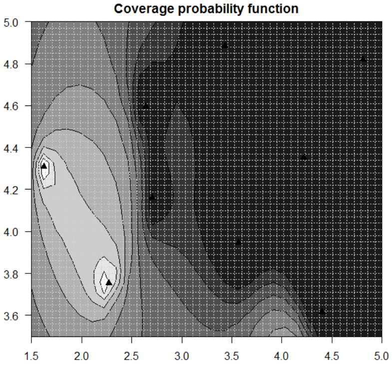
Under our Gaussian framework :

$$p_n(\mathbf{x}) = \phi\left(\frac{m_n(\mathbf{x}) - t}{\sqrt{k_n(\mathbf{x},\mathbf{x})}}\right),$$

with ϕ the Gaussian cumulative distribution function.

For any $\alpha \in [0,1]$, let us define

$$Q_\alpha = \{\mathbf{x} \in D \ : \ p_n(\mathbf{x}) \geq \alpha\}.$$

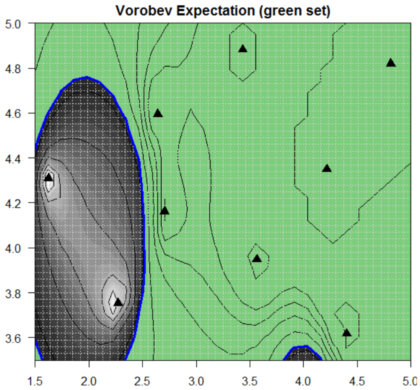


5. Ilya Molchanov (2006). *Theory of random sets.*
6. O Yu Vorob'ev (1984). *Srednemernoje modelirovanie (mean-measure modelling).*

Theory of random sets, Vorob'ev Expectation

The Vorob'ev Expectation of Γ is the set $Q_{\alpha_n^*}$ such that

$$\mu(Q_{\alpha_n^*}) = \mathbb{E}_n\left(\mu(\Gamma)\right) = \mathbb{E}\left(\mu(\Gamma) | Z(\mathcal{X}_n) = \mathbf{f}_n\right).$$



Vorob'ev deviation

$$\begin{aligned}\mathbb{E}_n[\mu(Q_{\alpha_n^*}\Delta\boldsymbol{\Gamma})] &= \mathbb{E}_n\left[\mu(\boldsymbol{\Gamma}\cap Q^c_{\alpha_n^*}) + \mu(Q_{\alpha_n^*}\cap\boldsymbol{\Gamma}^c)\right] \\ &= \int_{Q^c_{\alpha_n^*}} p_n(\mathbf{x})\mu(\mathrm{d}\mathbf{x}) + \int_{Q_{\alpha_n^*}} (1-p_n(\mathbf{x}))\mu(\mathrm{d}\mathbf{x})\end{aligned}$$

with $\mathbb{E}_n = \mathbb{E}[. \mid Z(\mathcal{X}_n) = \mathbf{f}_n]$.

Vorob'ev deviation

$$\mathbb{E}_n[\mu(Q_{\alpha_n^*}\Delta\boldsymbol{\Gamma})] = \mathbb{E}_n\left[\mu(\boldsymbol{\Gamma}\cap Q^c_{\alpha_n^*}) + \mu(Q_{\alpha_n^*}\cap\boldsymbol{\Gamma}^c)\right]$$
$$= \int_{Q^c_{\alpha_n^*}} p_n(\mathbf{x})\mu(\mathrm{d}\mathbf{x}) + \int_{Q_{\alpha_n^*}} (1-p_n(\mathbf{x}))\mu(\mathrm{d}\mathbf{x})$$

with $\mathbb{E}_n = \mathbb{E}[.\mid Z(\mathcal{X}_n) = \mathbf{f}_n]$.

The Vorob'ev expectation $Q_{\alpha_n^*}$ is the minimizer, among all closed subsets Q satisfying $\mu(Q) = \mathbb{E}_n[\mu(\boldsymbol{\Gamma})]$, of $\mathbb{E}_n[\mu(Q\Delta\boldsymbol{\Gamma})]$.

Infill strategy : *sampling strategy to choose new simulation points of f in order to reduce the Vorob'ev deviation.*

Stepwise Uncertainty Reduction (SUR strategy)

Uncertainty function : Vorob'ev deviation

$$\mathcal{H}_n^{\text{uncert}} = \mathbb{E}_n[\mu(Q_{\alpha_n^*} \Delta \mathbf{\Gamma})]$$
$$= \int_D \left(p_n(\mathbf{x}) \mathbb{1}_{\{p_n(\mathbf{x}) < \alpha_n^*\}} + (1 - p_n(\mathbf{x})) \mathbb{1}_{\{p_n(\mathbf{x}) \geq \alpha_n^*\}} \right) \mu(\mathrm{d}\mathbf{x})$$

Stepwise Uncertainty Reduction [7]

find the best next evaluation point $\mathbf{x}^{n+1}$ that optimally reduces the expected uncertainty $\mathcal{H}_{n+1}$ on the future estimate, i.e.,

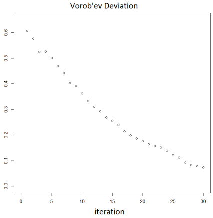
$$\mathbf{x}^{n+1} = \text{argmin}_{\mathbf{x} \in D} \mathbb{E}_{n,\mathbf{x}}[\mathcal{H}_{n+1}^{\text{uncert}}(\mathbf{x})].$$

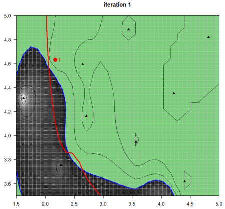
7. Julien Bect et al. (2012). ≪ Sequential design of computer experiments for the estimation of a probability of failure ≫. In : *Statistics and Computing* 22.3

An analytical example (3/5)

(left) Vorob'ev deviation

(right) boundary for the true Γ^* (red line) , estimation of Γ^* (green set)





On that application, the evaluations of f are computed analytically.

Our second objective is to include in the study the estimation of $f(\mathbf{x})$ from a low number of evaluations $g(\mathbf{x}, \mathbf{v}^j)$, $j = 1, \ldots, N$.

Table of Contents

An analytical example (4/5)

$$\Gamma^* = \{x \in D \ , \ f(x) \leq t\} \ , \text{ where } f(x) = \mathbb{E}[g(x, \mathbf{V})]$$

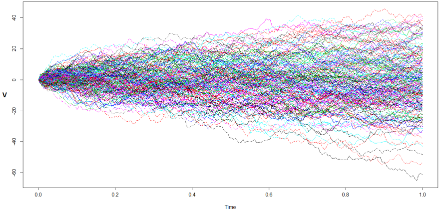


FIGURE – A finite sample of κ realizations of $\mathbf{V}$

We aim at estimating $\mathbb{E}[g(\mathbf{x}, \mathbf{V})]$ at points $\mathbf{x}^1, \ldots, \mathbf{x}^n$.

Crude scenario approach

The uncertainty of the functional variable $\mathbf{V}$ is represented by a subset of M curves $\mathbf{v}^1, \ldots, \mathbf{v}^M$ (randomly chosen among the initial sample of curves). Then we draw randomly with replacement N curves $\mathbf{w}^1, \ldots, \mathbf{w}^N$ in $\{\mathbf{v}^1, \ldots, \mathbf{v}^M\}$.

$$\Gamma^* = \{\mathbf{x} \in D \ , \ f(\mathbf{x}) = \mathbb{E}[g(\mathbf{x}, \mathbf{V})] \leq s\}$$
$$\simeq \{\mathbf{x} \in D \ , \ f(\mathbf{x}) \simeq \frac{1}{N} \sum_{i=1}^{N} g(\mathbf{x}, \mathbf{w}^i) \leq s\}.$$

Functional Principal Components

Functional PCA is a statistical method for investigating the dominant modes of variation of functional data.

Let $\mathcal{T}$ be a finite and closed interval of $\mathbb{R}$. Assume $\mathcal{F} \subset \mathbb{L}^2(\mathcal{T})$.

Then, $\mathbf{v}_1, \ldots, \mathbf{v}_\kappa$ can be viewed as independent realizations of a stochastic process $\mathbf{V}$ with unknown mean function $\mathbb{E}[V(t)] = \mu_{\mathbf{V}}(t)$ and covariance function $\text{Cov}\,(V(s), V(t)) = G_{\mathbf{V}}(s,t)$.

Then, if $G_{\mathbf{V}}$ is continuous on $\mathcal{T} \times \mathcal{T}$, it is well known that there exists an orthogonal expansion of $G_{\mathbf{V}}$ (in the $\mathbb{L}^2$ sense) in terms of eigenfunctions ϕ_k with associated eigenvalues λ_k (arranged in non increasing order), that is,

$$G_{\mathbf{V}}(s,t) = \sum_{k \geq 1} \lambda_k \phi_k(s) \phi_k(t) \,.$$

The random function $V(t)$, $t \in \mathcal{T}$ can be decomposed into an orthogonal expansion

$$V(t) = \mu_{\mathbf{V}}(t) + \sum_{k \geq 1} \sqrt{\lambda_k} \eta_k \phi_k(t)$$

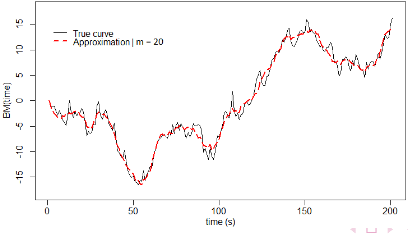
with η_k uncorrelated standardized random variables.

A truncation argument $V(t) = \mu_{\mathbf{V}}(t) + \sum_{k=1}^{m} \sqrt{\lambda_k} \eta_k \phi_k(t).$

Probability density estimation for the vector $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_m)^{\top}$
(non) parametric estimation.

Then, by sampling the vector $\boldsymbol{\eta}$, we get :

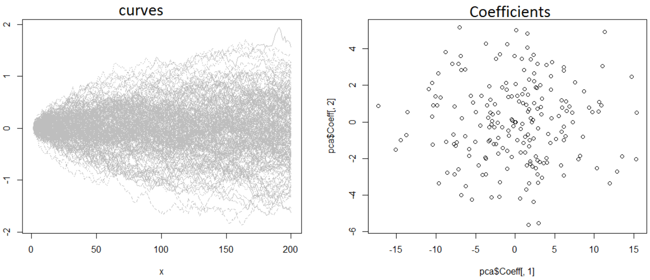
$$\Gamma^* = \{\mathbf{x} \in D \;,\; f(\mathbf{x}) = \mathbb{E}[g(\mathbf{x}, \mathbf{V})] \leq s\}$$
$$\simeq \{\mathbf{x} \in D \;,\; f(\mathbf{x}) \simeq \frac{1}{N} \sum_{i=1}^{N} g(\mathbf{x}, \boldsymbol{\eta}^i) \leq s\}.$$



A sequential hybrid methodology [8]

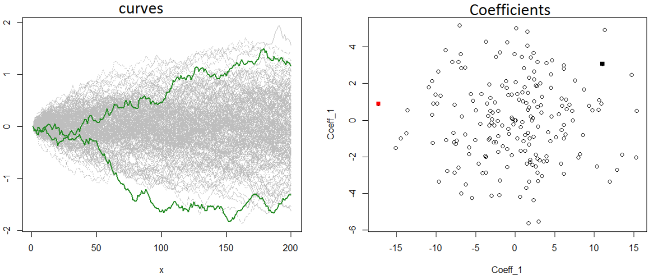
First Step : the adaptive selection of representative curves $\mathbf{v}^1, \ldots, \mathbf{v}^M$ among the initial sample of curves

- Let $\mathbf{v}^1, \ldots, \mathbf{v}^\kappa$ the initial sample of realizations of $\mathbf{V}$.
- We project the initial sample on the first m components in the functional PCA : $(\phi_k)_{k=1,\ldots,m} \rightarrow \boldsymbol{\eta}^1, \ldots, \boldsymbol{\eta}^\kappa$ κ points in $\mathbb{R}^m$.

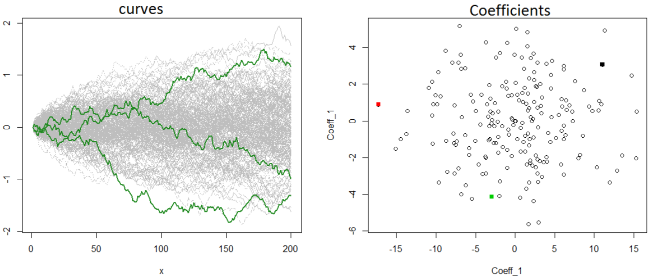


8. Mohamed Reda El Amri et al. (2018). ≪ Data-driven stochastic inversion under functional uncertainties ≫. URL : https://hal.inria.fr/hal-01704189

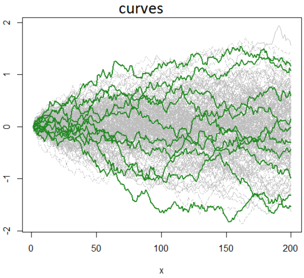
- We sequentially construct a space filling design in $\mathbb{R}^m$ by picking points in $\{\boldsymbol{\eta}^1, \ldots, \boldsymbol{\eta}^\kappa\}$.
- At each selected point $\boldsymbol{\eta}$ (on the right), we select the corresponding curve (on the left).

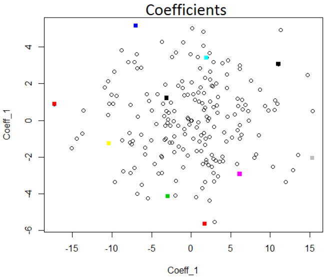


- The points η are chosen sequentially, with a min dist criterion in $\mathbb{R}^m$.



- We stop with N points $\eta^1, \ldots, \eta^N$ in $\mathbb{R}^m$ and the N corresponding curves $v^1, \ldots, v^N$ by considering a stopping criterion based on the the stability in the estimation of $\mathbb{E}[g(\mathbf{x}, \mathbf{V})]$.

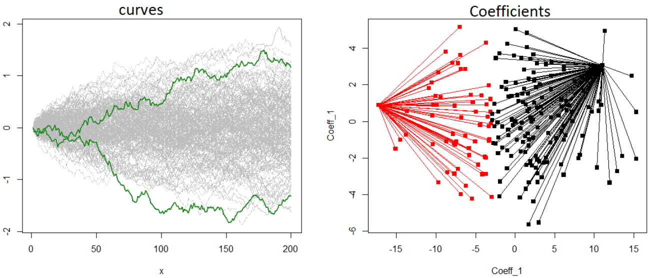


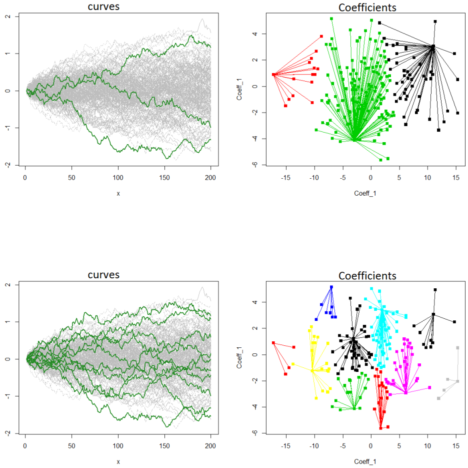


Second Step : the estimation of $\mathbb{E}[g(\mathbf{x},\mathbf{V})]$ with a weighted sum

- The computation of the weights is based on the Voronoi diagram.

$$\Gamma^* = \{x \in D \;,\; f(x) = \mathbb{E}[g(\mathbf{x},\mathbf{V})] \leq s\}$$
$$\simeq \{x \in D \;,\; f(\mathbf{x}) \simeq \frac{\sum_{i=1}^{N} w_i \; g(\mathbf{x},\mathbf{v}^i)}{\sum_{i=1}^{N} w_i} \leq s\}.$$




curves
Coefficients
Coeff_1
x
curves
Coefficients
Coeff_1
x

Global algorithm

Algorithm

Create an initial design of experiments (Doe) at n points in the control space $\mathbb{X}$
$l \leftarrow 2$
while Stopping criterion not met (SUR) **do**
  $x_{n+1} \leftarrow$ Using sampling criterion $\mathcal{J}_n$
  while Stopping criterion not met (Expectation Estimation) **do**
    Create design D_l and weight calculus w_l using (1)
    test $\leftarrow$ stopping criterion
    $l \leftarrow l + 1$
  end while
  Update Doe
  $n \leftarrow n + 1$
end while
end

Stopping criterion are based on the stability of the estimation of $\mathbb{E}[g(\mathbf{x}, \mathbf{V})]$ (hybrid weighted scenario approach) and on the stability of the Vorob'ev deviation (SUR strategy).

Analytical example (5/5)

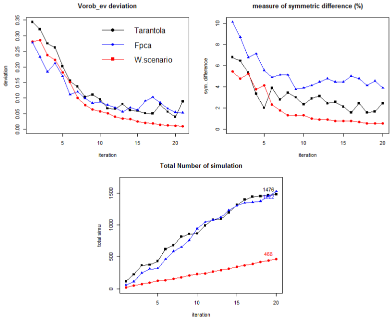
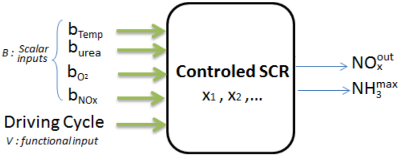


Figure – For Fpca and W.scenario, the truncation argument $m = 7$ is chosen to explain 97% of the variance of $\mathbf{V}$.

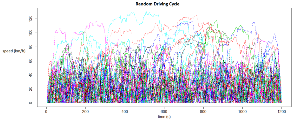
Test Case IFPEN : Selective Catalytic Reduction



- Thresholds :
  1. $\mathrm{NO_x^{out}} \leq 80\ mg.km^{-1}$
  2. $\mathrm{NH_3^{max}} \leq 30\ ppm$
- Objects of interest :
  1. $\{(\mathrm{x_1, x_2, ...}) : \mathbb{E}_{(\mathrm{B,V})}[\mathrm{NO_x^{out}}] \leq 80\ mg.km^{-1}\}$
  2. $\{(\mathrm{x_1, x_2, ...}) : \mathbb{E}_{(\mathrm{B,V})}[\mathrm{NH_3^{max}}] \leq 30\ ppm\}$
- Sensitivity Analysis : the driving cycle V is the most influential variable

IFPEN test case : control strategy for an automotive NO_x depollution system

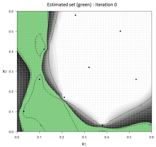
- $\mathbf{v}^1, \ldots, v^{100}$ random driving cycles.



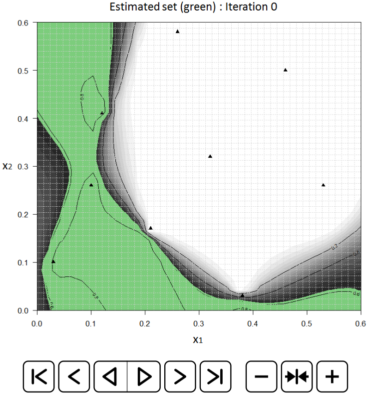
- $\mathrm{NH}_3^{\max}$ function of x_1 and x_2, $D = [0, 0.6]^2$

$$\Gamma^* = \{\mathbf{x} = (\mathrm{x}_1, \mathrm{x}_2) \in D \ : \ \mathbb{E}[\mathrm{NH}_3^{\max}(\mathbf{x}, \mathbf{V})] \leq s\}.$$

The model is expensive to evaluate (chemical kinetic model).



IFPEN test case, iteration 1, initial DoE : $n = 9$ black triangles


Estimated set (green) : Iteration 0
X2
X1
0.0
0.1
0.2
0.3
0.4
0.5
0.6

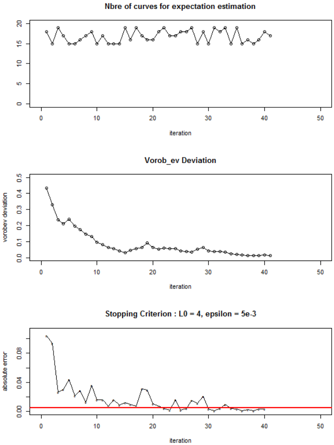

Nbre of curves for expectation estimation
iteration
Vorob_ev Deviation
vorobev deviation
iteration
Stopping Criterion : L0 = 4, epsilon = 5e-3
absolute error
iteration

Table of Contents

Conclusion

- Robust Inversion Problem :
  - Robustness measure : expectation ;
  - Sequential strategy SUR.
- With functional Uncertainty :
  - Hybrid sequential strategy based on functional PCA, a space filling strategy and a weighted scenario approach.

Perpectives

- Robustness measure : high probability, with some potentially small risk α

$$\Gamma_\alpha := \{\mathbf{x} \in D \ : \ \mathbb{P}\left(g(\mathbf{x}, \mathbf{V}) \leq s\right) \ \geq \ 1 - \alpha\} ;$$

- SUR strategy to choose both the next **x** and the next **v**.

Thanks for your attention !