Sensitivity analysis for optimization under constraints and with uncertainties

Kernel-based Sensitivity Analysis on (excursion) sets

Noé Fellmann
Céline Helbert & Christophette Blanchet (ECL)
Adrien Spagnol & Delphine Sinoquet (IFPEN)

École Centrale de Lyon & IFP Énergie nouvelles

Journées CIROQUO, 23-35 may 2023

From a complex system to robust optimization

Input parameters :
-component size
-material
-swell height
-...

inputs →

**Complex system :
Floating wind turbine**

output →

-Cost
-energy production
-environmental impact
-...

From a complex system to robust optimization

- The x are the deterministic inputs
- The u are uncertain inputs : $u = U(\omega)$ with U a random vector of density $\mathbb{P}_U$
- f is the objective function to minimize
- g is the constraint function defining the constraint to respect : $g \leq 0$

Robust optimization problem

$$x^* = \arg\min_{x \in \mathcal{X}} \mathbb{E}[f(x, U)] \text{ s.t. } \mathbb{P}[g(x, U) \leq 0] \geq P_{target}$$

How can Sensitivity Analysis be used to reduce the cost of a robust optimization ?

Robust Optimization and sensitivity analysis

Robust optimization problem...

$$x^* = \arg\min_{x \in \mathcal{X}} \mathbb{E}[f(x, U)] \text{ s.t. } \mathbb{P}[g(x, U) \leq 0] \geq P_{target}$$

... seen without the U

$$x^* = \arg\min_{x \in \mathcal{X}} F(x) \text{ s.t. } G(x) \leq 0$$

- Goal Oriented Sensitivity Analysis to reduce the input dimension : Spagnol 2020
- Huge evaluation cost of F and G
- What about the U ?

How to quantify the impact of the uncertain inputs U on the optimization ?

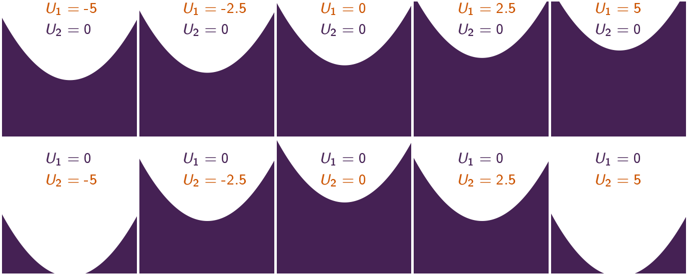
$\rightarrow$ SA on excursion sets : $\mathrm{U} \overset{\phi}{\mapsto} \Gamma = \{x \in \mathcal{X}, g(x, \mathrm{U}) \leq 0\}$

A toy excursion set

Toy function from [El-Amri et al. 2021]

$$\forall (x, u) \in [-5, 5]^4 \;\; g(\boldsymbol{x}, \boldsymbol{u}) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

$$\Gamma_u = \{x \in [-5, 5]^2, g(x, u) \leq 0\}$$



SA on excursion sets

Excursion sets

New output :

$$\Gamma_U = \{x \in \mathcal{X}, g(x, U) \leq 0\}, \tag{1}$$

which is called a random excursion set.

Influence of the uncertain inputs U on Γ_U ? $\Rightarrow$ SA on excursion sets.
Yet most SA techniques have scalar or vectorial outputs.
How then can we do sensitivity analysis on (excursion) sets ?

– SA on sets using random set theory notably Vorob'ev expectation and deviation
– SA on sets using universal indices from Fort, Klein et Lagnoux 2021
– **SA on sets using RKHS theory :**
$\rightarrow$ one article (pre-print on HAL)

} One applicative article planned

Table of Contents

Table of Contents

Distribution embedding into a RKHS

Dependence measures : $S_i = ||\mathbb{P}_{(X_i,Y)} - \mathbb{P}_{X_i} \otimes \mathbb{P}_Y||$

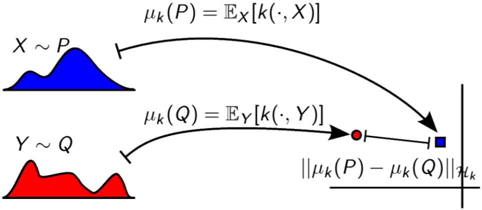


Figure – Kernel mean embedding

with k a (positive definite) kernel $k : (x, x') \in \mathcal{X}^2 \mapsto k(x, x') \in \mathbb{R}$.

Hilbert Schmidt Independence Criterion (HSIC), Gretton et al. 2006

With $K = k_{\mathcal{X}_i} \otimes k_{\mathcal{Y}}$, the HSIC is given by :

$$\mathrm{HSIC}_K(X_i, Y) = ||\mu_K(X_i, Y) - \mu_{k_{\mathcal{X}_i}}(X_i) \otimes \mu_{k_{\mathcal{Y}}}(Y)||^2_{\mathcal{H}_K}$$

When K is characteristic (injectivity of the mean embedding),

$$\mathrm{HSIC}_K(X_i, Y) = 0 \text{ iif } X_i \perp Y \rightarrow \text{ screening.}$$

HSIC-ANOVA index [daVeiga 2021]

Assuming that the inputs are independent and that the input kernels are ANOVA,

$$\mathrm{HSIC}(\boldsymbol{X}, Y) = \sum_{A \subseteq \{1,\dots,d\}} \sum_{B \subseteq A} (-1)^{|A|-|B|} \, \mathrm{HSIC}\left(\boldsymbol{X}_B, Y\right).$$

HSIC-ANOVA indices are then defined as :

$$S_i^{\mathrm{HSIC}} := \frac{\mathrm{HSIC}(X_i, Y)}{\mathrm{HSIC}(\boldsymbol{X}, Y)},$$

$$S_{T_i}^{\mathrm{HSIC}} := 1 - \frac{\mathrm{HSIC}(\boldsymbol{X}_{-i}, Y)}{\mathrm{HSIC}(\boldsymbol{X}, Y)}$$

and are suited for ranking (and screening).

- Easy to estimate :

$$\begin{aligned}\mathrm{HSIC}_{k_{\mathcal{X}}, k_{\mathcal{Y}}}(X, Y) = \ & \mathbb{E}[k_{\mathcal{X}}(X, X') k_{\mathcal{Y}}(Y, Y')] \\ & + \mathbb{E}[k_{\mathcal{X}}(X, X')] \mathbb{E}[k_{\mathcal{Y}}(Y, Y')] \\ & - 2\mathbb{E}[\mathbb{E}[k_{\mathcal{X}}(X, X')|X] \mathbb{E}[k_{\mathcal{Y}}(Y, Y')|Y]].\end{aligned}$$

- Only requirement : to have kernels on the inputs and on the output

Table of Contents

SA on sets : a kernel between sets

With $A\Delta B = A \cup B - B \cap A$ and λ the Lebesgue measure, we define a kernel on the Lebesgue σ-algebra $\mathcal{B}(\mathcal{X})$ by :

$$\forall \Gamma_1, \Gamma_2 \in \mathcal{B}(\mathcal{X}),\ k_{set}(\Gamma_1, \Gamma_2) = \exp\left(-\frac{\lambda(\Gamma_1 \Delta \Gamma_2)}{2\sigma^2}\right).$$

Proposition (A kernel between sets)

k_{set} *is a kernel* [Balança et Herbin 2012] *and is characteristic.*

k_{set} is characteristic, sketch of proof

- $\mathcal{B}(\mathcal{X}) \to \mathcal{B} = \mathcal{B}(\mathcal{X})/\sim_\delta$ where δ is the volume of the symmetric difference and $\sim_\delta$ the equivalent relation $A \sim_\delta B$ iif $\delta(A, B) = 0$ i.e. A and B are equal except on a λ-negligible set.
- We show that $(\mathcal{B}, \delta)$ is a Polish space (separable completely metrizable topological space). $(\mathcal{B}, \delta)$ is a metric space so we just need separability and completeness. Separability holds as "it's a subspace of $L_2(\mathcal{X})$" and we show that it's closed.
- We use a Proposition from Ziegel, Ginsbourger et Dümbgen 2022,

Proposition

Let $\mathcal{B}$ be a Polish space, H a separable Hilbert space, T a measurable and injective mapping from $\mathcal{B}$ to H, and $\varphi \in \Phi_\infty^+$. Then, the kernel k on $\mathcal{B}$ defined by

$$k\left(\gamma, \gamma'\right) := \varphi\left(\left\|T(\gamma) - T\left(\gamma'\right)\right\|_H^2\right), \quad \gamma, \gamma' \in \mathcal{B}$$

is integrally strictly positive definite with respect to $\mathcal{M}(\mathcal{B})$ (which implies that it is characteristic).

with $H = L_2(\mathcal{X})$, $\varphi = \exp(-\frac{\cdot}{2\sigma^2})$ and T defined by $T(\gamma) := x \mapsto \mathbb{1}_\gamma(x)$ for any $\gamma \in \mathcal{B}$ so that $\|T(\gamma) - T\left(\gamma'\right)\|_H^2 = \lambda(\gamma \Delta \gamma')$.

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_l, \Gamma) := \mathrm{HSIC}_{k_l, k_{set}}(U_l, \Gamma) = \mathbb{E}\left[(k_l(U_l, U_l{}') - 1)k_{set}(\Gamma, \Gamma')\right]$

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_I, \Gamma) := \mathrm{HSIC}_{k_I, k_{set}}(U_I, \Gamma) = \mathbb{E}\left[(k_I(U_I, U_I') - 1) k_{set}(\Gamma, \Gamma')\right]$
- $\widehat{\mathrm{H}_{set}}\,(U_I, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_I\left(U_I^{(i)}, U_I^{(j)}\right) - 1\right) k_{set}\left(\Gamma^{(i)}, \Gamma^{(j)}\right)$

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_I, \Gamma) := \mathrm{HSIC}_{k_I, k_{set}}(U_I, \Gamma) = \mathbb{E}\left[(k_I(U_I, U_I{}') - 1)k_{set}(\Gamma, \Gamma')\right]$
- $\widehat{\mathrm{H}_{set}}\,(U_I, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_I\left(U_I^{(i)}, U_I^{(j)}\right) - 1\right) k_{set}\left(\Gamma^{(i)}, \Gamma^{(j)}\right)$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X_{i,j}^{(k)})) \rightarrow n(n-1)m$ tests of $X \in \Gamma$

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_l, \Gamma) := \mathrm{HSIC}_{k_l, k_{set}}(U_l, \Gamma) = \mathbb{E}\left[(k_l(U_l, U_l{}') - 1) k_{set}(\Gamma, \Gamma')\right]$
- $\widehat{\mathrm{H}_{set}}\,(U_l, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_l \left(U_l^{(i)}, U_l^{(j)} \right) - 1 \right) k_{set}\, \left(\Gamma^{(i)}, \Gamma^{(j)} \right)$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X_{i,j}^{(k)})) \rightarrow$ $n(n-1)m$ tests of $X \in \Gamma$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X^{(k)}))$

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_I,\Gamma) := \mathrm{HSIC}_{k_I,k_{set}}(U_I,\Gamma) = \mathbb{E}\left[(k_I(U_I,U_I') - 1)k_{set}(\Gamma,\Gamma')\right]$
- $\widehat{\mathrm{H}_{set}}(U_I,\Gamma) = \frac{2}{n(n-1)}\sum_{i<j}^{n}\left(k_I\left(U_I^{(i)},U_I^{(j)}\right) - 1\right)k_{set}\left(\Gamma^{(i)},\Gamma^{(j)}\right)$
- $\widehat{k_{set}}(\Gamma^{(i)},\Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2}\frac{1}{m}\sum_{k=1}^{m}\mathbb{1}_{\Gamma^{(i)}\Delta\Gamma^{(j)}}(X_{i,j}^{(k)}))\rightarrow n(n-1)m$ tests of $X \in \Gamma$
- $\widehat{k_{set}}(\Gamma^{(i)},\Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2}\frac{1}{m}\sum_{k=1}^{m}\mathbb{1}_{\Gamma^{(i)}\Delta\Gamma^{(j)}}(X^{(k)}))$
- $\widehat{\widehat{\mathrm{H}_{set}}}(U_I,\Gamma) = \frac{2}{n(n-1)}\sum_{i<j}^{n}\left(k_I\left(U_I^{(i)},U_I^{(j)}\right) - 1\right)\exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2}\frac{1}{m}\sum_{k=1}^{m}\mathbb{1}_{\Gamma^{(i)}\Delta\Gamma^{(j)}}(X^{(k)}))$

HSIC-ANOVA on sets, estimation

- $H_{set}(U_I, \Gamma) := \mathrm{HSIC}_{k_I, k_{set}}(U_I, \Gamma) = \mathbb{E}\left[(k_I(U_I, U_I') - 1) k_{set}(\Gamma, \Gamma')\right]$
- $\widehat{H_{set}}(U_I, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_I\left(U_I^{(i)}, U_I^{(j)}\right) - 1\right) k_{set}\left(\Gamma^{(i)}, \Gamma^{(j)}\right)$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp\left(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X_{i,j}^{(k)})\right) \rightarrow n(n-1)m$ tests of $X \in \Gamma$
- $\widehat{k_{set}}(\Gamma^{(i)}, \Gamma^{(j)}) = \exp\left(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X^{(k)})\right)$
- $\widehat{\widehat{H_{set}}}(U_I, \Gamma) = \frac{2}{n(n-1)} \sum_{i<j}^{n} \left(k_I\left(U_I^{(i)}, U_I^{(j)}\right) - 1\right) \exp\left(-\frac{\lambda(\mathcal{X})}{2\sigma^2} \frac{1}{m} \sum_{k=1}^{m} \mathbb{1}_{\Gamma^{(i)} \Delta \Gamma^{(j)}}(X^{(k)})\right)$

Proposition

The quadratic risk of the nested estimator $\widehat{\widehat{H_{set}}}$ *verifies :*

$$\mathbb{E}\left(\widehat{\widehat{H_{set}}}(U_I, \Gamma) - H_{set}(U_I, \Gamma)\right)^2 \leq 2\left(\frac{2\sigma_1^2}{n(n-1)} + \frac{4(n-2)\sigma_2^2}{n(n-1)} + \frac{L^2\sigma_3^2}{m}\right).$$

HSIC-ANOVA on sets, estimation

- $\mathrm{H}_{set}(U_I,\Gamma) := \mathrm{HSIC}_{k_I,k_{set}}(U_I,\Gamma) = \mathbb{E}\left[(k_I(U_I,U_I{}') - 1)k_{set}(\Gamma,\Gamma')\right]$
- $\widehat{\mathrm{H}_{set}}(U_I,\Gamma) = \frac{2}{n(n-1)}\sum_{i<j}^{n}\left(k_I\left(U_I^{(i)},U_I^{(j)}\right) - 1\right)k_{set}\left(\Gamma^{(i)},\Gamma^{(j)}\right)$
- $\widehat{k_{set}}(\Gamma^{(i)},\Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2}\frac{1}{m}\sum_{k=1}^{m}\mathbb{1}_{\Gamma^{(i)}\Delta\Gamma^{(j)}}(X_{i,j}^{(k)}))\rightarrow n(n-1)m$ tests of $X \in \Gamma$
- $\widehat{k_{set}}(\Gamma^{(i)},\Gamma^{(j)}) = \exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2}\frac{1}{m}\sum_{k=1}^{m}\mathbb{1}_{\Gamma^{(i)}\Delta\Gamma^{(j)}}(X^{(k)}))$
- $\widehat{\widehat{\mathrm{H}_{set}}}(U_I,\Gamma) = \frac{2}{n(n-1)}\sum_{i<j}^{n}\left(k_I\left(U_I^{(i)},U_I^{(j)}\right) - 1\right)\exp(-\frac{\lambda(\mathcal{X})}{2\sigma^2}\frac{1}{m}\sum_{k=1}^{m}\mathbb{1}_{\Gamma^{(i)}\Delta\Gamma^{(j)}}(X^{(k)}))$

Proposition

The quadratic risk of the nested estimator $\widehat{\widehat{\mathrm{H}_{set}}}$ *verifies :*

$$\mathbb{E}\left(\widehat{\widehat{\mathrm{H}_{set}}}(U_I,\Gamma) - \mathrm{H}_{set}(U_I,\Gamma)\right)^2 \leq 2\left(\frac{2\sigma_1^2}{n(n-1)} + \frac{4(n-2)\sigma_2^2}{n(n-1)} + \frac{L^2\sigma_3^2}{m}\right).$$

We can now compute $S_i^{\widehat{\widehat{\mathrm{H}_{set}}}}$ or $S_{T_i}^{\widehat{\widehat{\mathrm{H}_{set}}}}$ to perform SA on set-valued outputs.

Table of Contents

Toy function 1

From El-Amri et al. 2021,

$$\forall x, u \in [-5,5]^2 \times [-5,5]^3 \quad g(x_1, x_2, u_1, u_2, u_3) = -x_1^2 + 5x_2 - u_1 + u_2^2 - 1$$

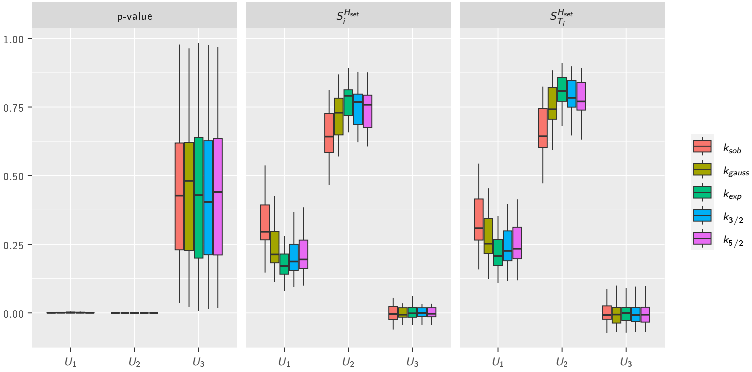


Figure – Estimation of the p-values, $\hat{\hat{S}}_i^{H_{set}}$ and $\hat{\hat{S}}_{T_i}^{H_{set}}$ for the excursion set defined by the constraint $g \leq 0$ computed for 5 input kernels with $n = 100$, $m = 100$ and repeated 20 times

Toy function 1, quadratic risk

Proposition

The quadratic risk of the nested estimator $\widehat{\mathrm{H}_{set}}$ verifies :

$$\mathbb{E}\left(\widehat{\mathrm{H}_{set}}(U_l,\Gamma) - \mathrm{H}_{set}(U_l,\Gamma)\right)^2 \leq 2\left(\frac{2\sigma_1^2}{n(n-1)} + \frac{4(n-2)\sigma_2^2}{n(n-1)} + \frac{L^2\sigma_3^2}{m}\right).$$

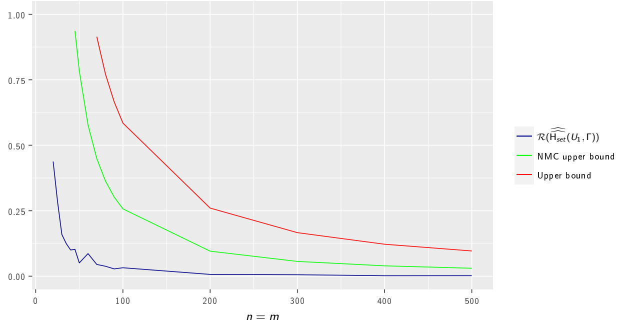


Figure – Evolution of $\mathcal{R}(\widehat{\mathrm{H}_{set}}(U_1,\Gamma))$ and of the associated upper bounds for the excursion set of the constraint $g \leq 0$

Pollutant concentration maps : Maps of Sobol indices

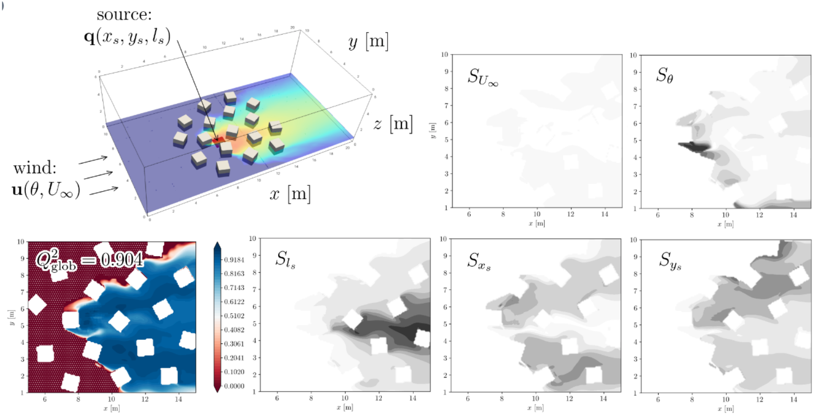


Figure – Maps of the sobol indices of pollutant dispersion (Mathis Pasquier)

Interpretation :

$$S_{l_s} \gg S_{x_s} \approx S_{y_s} \approx S_\theta \gg S_{U_\infty}$$

Kernel-based SA on pollutant concentration maps

Sobol map interpretation : $S_{l_s} \gg S_{x_s} \approx S_{y_s} \approx S_\theta \gg S_{U_\infty}$.
$\forall(x,y) \in [5,15] \times [1,10]$, $g(x,y,U)$ is the pollutant concentration at the point (x,y) for a given uncertain parameter U. What is the set-valued output ?

- Test 1 : $\Gamma_U = \{(x,y) \in [5,15] \times [1,10], g(x,y,U) \geq C_{seuil}\}$. C_{seuil} to choose (toxicity threshold).

	θ	U_∞	x_s	y_s	l_s
P-value	$6.6.10^{-4}$	0.11	0	0	0
$S_i^{H_{set}}$	0.069	0.016	0.25	0.15	0.48

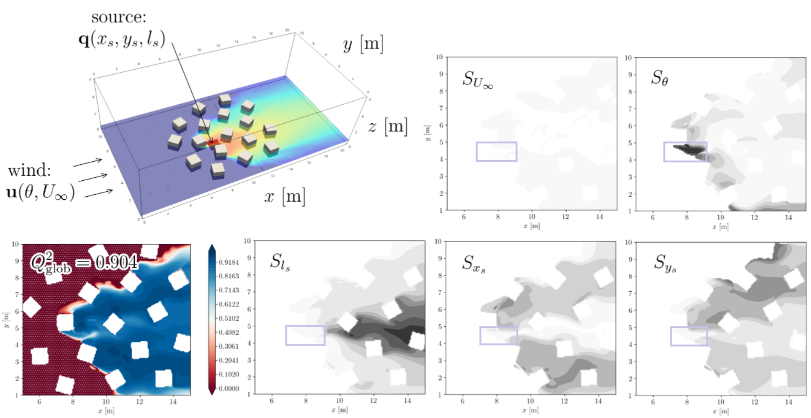
- Test 2 : $\Gamma_U = \{(x,y,z) \in [5,15] \times [1,10] \times [C_{min}, C_{max}], z \leq g(x,y,U)\}$

	θ	U_∞	x_s	y_s	l_s
P-value	$6.6.10^{-3}$	0.77	0.026	0.010	0
$S_i^{H_{set}}$	0.11	0.0081	0.20	0.17	0.50

In both cases we obtain :

$$S_{l_s} > S_{x_s} > S_{y_s} > S_\theta > S_{U_\infty}$$

Kernel-based SA on pollutant concentration maps : subspace



$$\Gamma_U = \{(x, y, C) \in [7, 9] \times [4, 5] \times [C_{min}, C_{max}], C \leq g(x, y, U)\}$$

	θ	U_∞	x_s	y_s	l_s
P-value	0	0.002	0	0	0.03
$S_i^{H_{set}}$	0.59	0.09	0.14	0.13	0.04

Outlook of the thesis

- Does this work in high dimensional space $\mathcal{X}$?
  - Very small volumes ? $\rightarrow$ importance sampling ?
  - If we know that the "effective" dimension d' of Γ_U is small, can we still conduct the sensitivity analysis on $\mathcal{X}$ or should we reduce the dimension on the x before ?
- How / where to use a sensitivity analysis on the U (and on the x) inside a robust optimization methodology ?

El-Amri, Reda et al. (2021). *A sampling criterion for constrained Bayesian optimization with uncertainties*. arXiv : 2103.05706 [stat.ML].

Balança, Paul et Erick Herbin (jan. 2012). "A set-indexed Ornstein-Uhlenbeck process". In : *Electronic Communications in Probability* 17.none. doi : 10.1214/ecp.v17-1903. url : https://doi.org/10.1214%2Fecp.v17-1903.

Cousin, Alexis (oct. 2021). "Chance constraint optimization of a complex system : Application to the design of a floating offshore wind turbine". Theses. Institut Polytechnique de Paris. url : https://tel.archives-ouvertes.fr/tel-03500604.

daVeiga, Sébastien (jan. 2021). "Kernel-based ANOVA decomposition and Shapley effects - Application to global sensitivity analysis". working paper or preprint. url : https://hal.archives-ouvertes.fr/hal-03108628.

Fellmann, Noé et al. (mai 2023). "Kernel-based sensivity analysis for (excursion) sets". working paper or preprint. url : https://hal-ifp.archives-ouvertes.fr/hal-04094561.

Fort, Jean-Claude, Thierry Klein et Agnès Lagnoux (jan. 2021). "Global Sensitivity Analysis and Wasserstein Spaces". In : *SIAM/ASA Journal on Uncertainty Quantification* 9.2. Publisher : Society for Industrial and Applied Mathematics, p. 880-921. doi : 10.1137/20M1354957. url : https://epubs.siam.org/doi/abs/10.1137/20M1354957.

Gretton, Arthur et al. (2006). “A Kernel Method for the Two-Sample-Problem”. In : *Advances in Neural Information Processing Systems*. T. 19. MIT Press. url : https://proceedings.neurips.cc/paper/2006/hash/e9fb2eda3d9c55a0d89c98d6c54b5b3e-Abstract.html.

Spagnol, Adrien (juill. 2020). “Kernel-based sensitivity indices for high-dimensional optimization problems”. Theses. Université de Lyon. url : https://tel.archives-ouvertes.fr/tel-03173192.

Ziegel, Johanna, David Ginsbourger et Lutz Dümbgen (2022). *Characteristic kernels on Hilbert spaces, Banach spaces, and on sets of measures*. arXiv : 2206.07588 [stat.ML].