Gradient-Based Active Learning with Gaussian Processes for Global Sensitivity Analysis

Application to a water and pesticide transport model

MASCOT-NUM 2026

Guerlain Lambert[1,2] Céline Helbert[1] Claire Lauvernet[2]

[1]Centrale Lyon, Institut Camille Jordan, Écully

[2]INRAE Lyon, RiverLy, Villeurbanne

April, 1st 2026

Motivation : Wine

What is the connection between this bottle of Beaujolais and UQ ?

(De-)Motivation : Welcome to Pesticidstan


Le Monde

FR

Guerres au Proche-Orient · Élections municipales 2026 · International · Planète · Politique · Société · Économie · Idées · Culture · Le Goût du Monde

SCIENCES • BIODIVERSITÉ

L'usage de pesticides bouleverserait en profondeur la biodiversité microbienne des sols

Le Monde

Guerres au Proche-Orient | Élections municipales 2026 | International | Planète | Politique | Société | Économie

SCIENCES • BIODIVERSITÉ

L'usage de pesticides bouleversera profondeur la biodiversité microbien

Courrier international

La France vue de l'étranger | Société | Économie | Sciences et environnement | Culture | RÉVEIL | STORIES | VIDÉOS

SCIENCES ET ENVIRONNEMENT • PESTICIDES

Pesticides. L'exposition au chlorpyrifos accélère le vieillissement de poissons d'eau douce


Le Monde
Politique
Société
Économie
Guerres au Proche-Orient
Élections municipales 2026
SCIENCES • BIODIVERSITÉ
Courrier international
La France vue de l'étranger
Société
Économie
Sciences et environnement
Culture
RÉVEIL
STORIES
VIDÉOS
SCIENCES ET ENVIRONNEMENT • PESTICIDES
Le Monde
Menu
Planète
Comprendre le réchauffement climatique
9 indicateurs de l'urgence climatique
PLANÈTE • BIODIVERSITÉ
Là où les pesticides sont le plus vendus, les oiseaux reculent

(De-)Motivation : Welcome to Pesticidstan

(De-)Motivation : Welcome to Pesticidstan

(De-)Motivation : Welcome to Pesticidstan

Motivation

This poses a public health and environmental issue : it needs to be modelized and analyzed.

Modeling **pesticide** transport is based on the larger model **CATHY** (CATchment HYdrology) [2, 9]. It offers :

Motivation

This poses a public health and environmental issue : it needs to be modelized and analyzed.

Modeling **pesticide** transport is based on the larger model **CATHY** (CATchment HYdrology) [2, 9]. It offers :

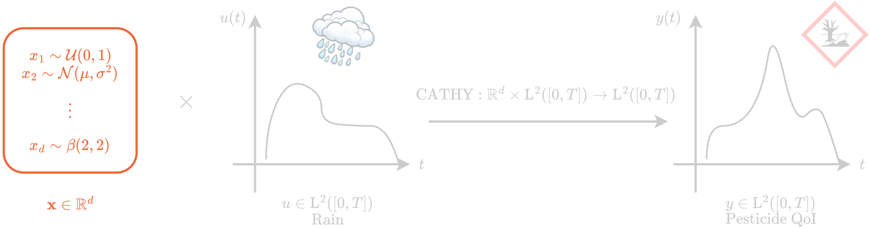
- a 100% physics-based model solving coupled surface/subsurface flow via the 3D Richards P.D.E.:

$$S_w S_s \frac{\partial \psi}{\partial t} + \phi \frac{\partial S_w}{\partial t} = \nabla [K_s K_r (\nabla \psi + \eta_z)] + q_{ss}$$

- a simulations of coupled surface and subsurface hydrological processes involving water flow and solute transport, including advection–dispersion:

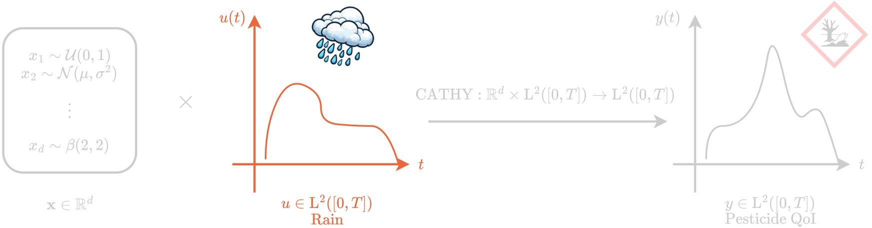
$$\frac{\partial C}{\partial t} = \nabla (D \nabla c) - \nabla (\vec{V} c) + R$$

Figure: CATHY is constructed on the entire catchment.

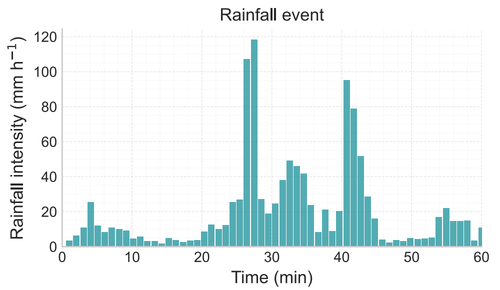


- Ksat : Saturated hydraulic conductivity (m/s)
- θ_s : Saturated water content (-).
- α_{VG} : Air-entry pressure inverse (m^{-1})
- K_d : Distribution coefficient (L/kg)
- Initial groundwater depth, Strickler roughness (rugosity), ...

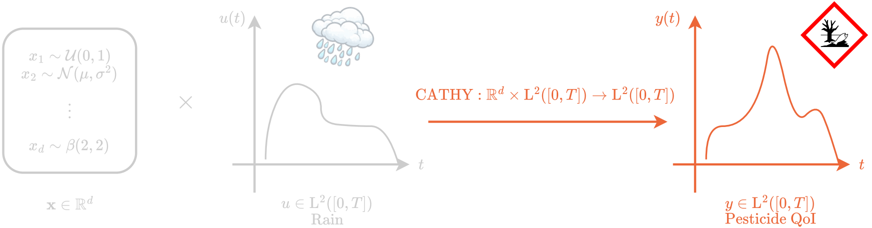
Motivation : model setup



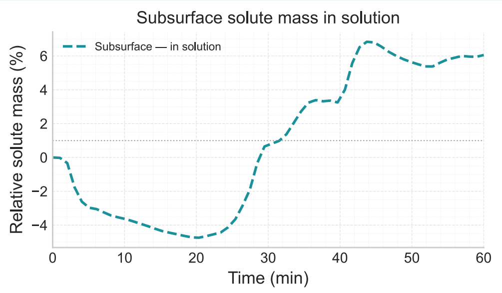
- Study setting : INRAE experimental site in Beaujolais (Morcille)
- Rainfall events of 1 hour.
- Representative of Beaujolais rainfalls.
- Usually fixed in hydrology GSA studies.
- Here, considered as a variable.



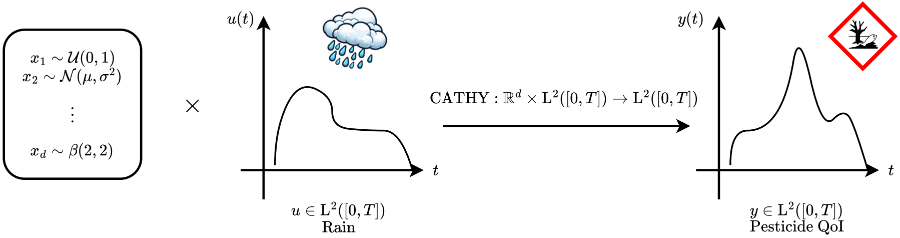
Motivation : model setup



- Negative values: pesticide still at the surface or adsorbed — not yet mobilised.
- Positive values (flushing): infiltrating water delivers a concentrated pulse of solute to the subsurface.
- Pollution risk: this first flush feeds baseflow at the outlet and contaminates groundwater resources.



Motivation : model setup



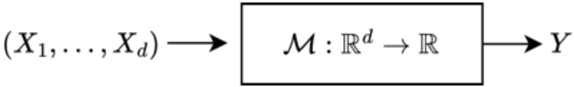
- Costly-to-evaluate model
- Large input space (spatio-temporal data, soil, pesticide parameters, ...).
- Rain as a variable of the model.

Problematic

Under constrained budget allocation, tailor an active learning strategy in a context of:

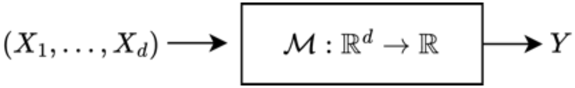
1. Goal oriented for GSA to understand physical processes and assess the Rain influence.
2. Building a *good* surrogate model.

Recall on variance-based sensitivity measures



How do you measure how variables affect the variability of a model's output?

Recall on variance-based sensitivity measures



How do you measure how variables affect the variability of a model's output?

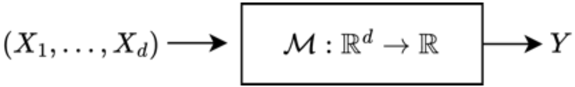
- **First-Order Sobol Index** ($S_{\mathbf{A}}$):

$$S_{\mathbf{A}} = \frac{\mathrm{Var}\left(\mathbb{E}\left[\mathcal{M}(\mathbf{X}) \mid \mathbf{X}_{\mathbf{A}}\right]\right)}{\mathrm{Var}(\mathcal{M}(\mathbf{X}))}$$

- **Total Sobol Index** ($S^T_{\mathbf{A}}$):

$$S^T_{\mathbf{A}} = 1 - S_{\sim\mathbf{A}}$$

Recall on variance-based sensitivity measures



How do you measure how variables affect the variability of a model's output?

- **First-Order Sobol Index** $(S_{\mathbf{A}})$:

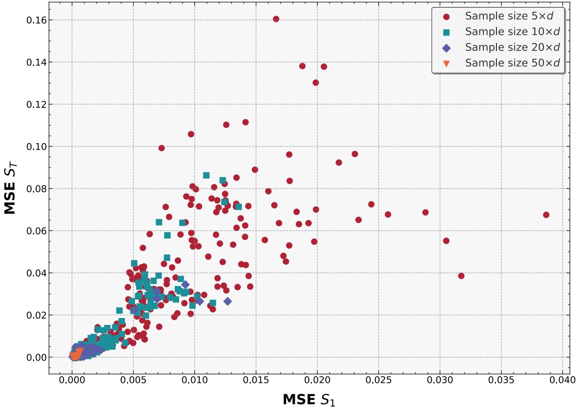
$$S_{\mathbf{A}} = \frac{\mathrm{Var}\left(\mathbb{E}\left[\mathcal{M}(\mathbf{X}) \mid \mathbf{X}_{\mathbf{A}}\right]\right)}{\mathrm{Var}(\mathcal{M}(\mathbf{X}))}$$

- **Total Sobol Index** $(S^T_{\mathbf{A}})$:

$$S^T_{\mathbf{A}} = 1 - S_{\sim\mathbf{A}}$$

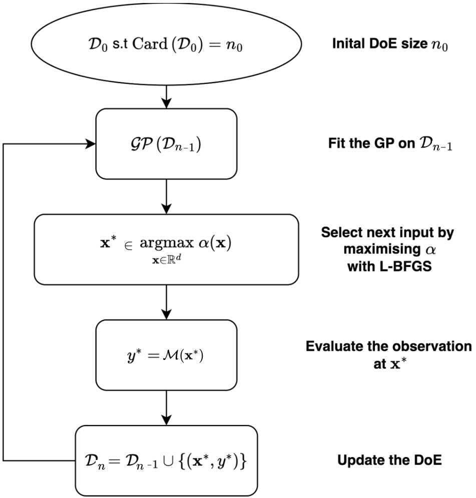
Does some DoE for surrogate fitting yield to better Sobol indices estimation ?

Recall on main sensitivity indices



Oui.

Active Learning strategy



Stopping criterion: Total budget.
Acquisition function: Evaluation on GP posterior.

How can we define an acquisition function that would provide the *most* information for estimating Sobol?

Derivative-based sensitivity measures

Find an alternative to variance-based sensitivity indices.

Derivative-based sensitivity measures

Find an alternative to variance-based sensitivity indices.

Let $\mathbf{X} \in \mathrm{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ a random vector of $\mathbb{R}^d$. The k^{th} DGSM :

$$D_k = \mathbb{E}_X \left[\left(\partial_k \mathcal{M}(\mathbf{X}) \right)^2 \right]$$

Derivative-based sensitivity measures

Find an alternative to variance-based sensitivity indices.

Let $\mathbf{X} \in \mathrm{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ a random vector of $\mathbb{R}^d$. The k^{th} DGSM :

$$D_k = \mathbb{E}_X \left[\left(\partial_k \mathcal{M}(\mathbf{X}) \right)^2 \right]$$

Can we bridge the gap between Variance-based and Derivative-based measures ?

Derivative-based sensitivity measures

Find an alternative to variance-based sensitivity indices.

Let $\mathbf{X} \in \mathrm{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ a random vector of $\mathbb{R}^d$. The k^{th} DGSM :

$$D_k = \mathbb{E}_X \left[\left(\partial_k \mathcal{M}(\mathbf{X}) \right)^2 \right]$$

Can we bridge the gap between Variance-based and Derivative-based measures ?

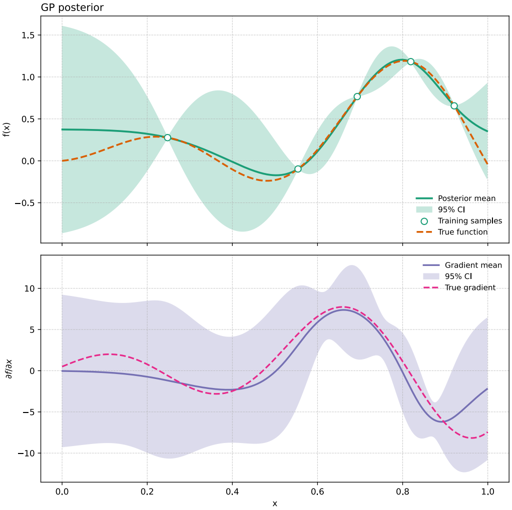
Under mild assumptions,

$$S_k^T \leq \frac{C_{X_k} D_k}{\mathrm{Var}(\mathcal{M}(\mathbf{X}))}$$

with, C_{X_k} the Poincaré constant.

References : [8, 7, 3]

Gaussian Process regression and its gradient



DGSM estimation via GP : [6]

Derivative of Gaussian process

$$\begin{bmatrix} \boldsymbol{\eta} \\ \partial_i \eta(\mathbf{x}_*) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{m} \\ \partial_i m_* \end{bmatrix}, \begin{bmatrix} \mathbf{K} & \partial_i \mathbf{k}_* \\ (\partial_i \mathbf{k}_*)^\top & \partial_i^2 k_{**} \end{bmatrix} \right)$$

$$\partial_i \eta(\mathbf{x}_*) \mid \mathcal{D} \sim \mathcal{N}\left(\mu'_{*,i}, (\sigma^2_{*,i})' \right)$$

$$\mu'_{*,i} = \partial_i m_* + (\partial_i \mathbf{k})^\top \mathbf{K}^{-1} (y - \mathbf{m})$$

$$(\sigma^2_{*,i})' = \partial_i^2 k_{**} - (\partial_i \mathbf{k})^\top \mathbf{K}^{-1} \partial_i \mathbf{k}_*$$

For popular kernels (ARD-RBF, ARD-Matérn, ...), the gradient is known in closed form and/or can benefit of auto-differentiation.

Derivative-based Active Learning : existing work

$$\frac{(\partial_i \eta(\mathbf{x}))^2}{\sigma_i'(\mathbf{x})^2} \mid \mathcal{D} \sim \chi_1^2\left(\frac{\mu_i'(\mathbf{x})^2}{\sigma_i'(\mathbf{x})^2}\right)$$

Hence,

$$\sigma_{i,sq}^2(\mathbf{x}) = 4\sigma_i'(\mathbf{x})^2\mu_i'(\mathbf{x}) + 2\sigma_i'(\mathbf{x})^4$$

Acquisition functions [1]

Maximum Variance :

$$\alpha_{\mathrm{Var}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2$$

Variance Reduction :

$$\alpha_{\mathrm{VarRed}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2 - \sigma_{i,sq}^{\ell}(\mathbf{x})^2$$

Entropy :

$$\alpha_{\mathrm{Ent}}(\mathbf{x}) = \sum_{i=1}^{d} H_{i,sq}(\mathbf{x}) - H_{i,sq}^{\ell}(\mathbf{x})$$

Derivative-based Active Learning : existing work

Acquisition functions [1]

Maximum Variance :

$$\alpha_{\mathrm{Var}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2$$

Variance Reduction :

$$\alpha_{\mathrm{VarRed}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2 - \sigma_{i,sq}^{\ell}(\mathbf{x})^2$$

Entropy :

$$\alpha_{\mathrm{Ent}}(\mathbf{x}) = \sum_{i=1}^{d} H_{i,sq}(\mathbf{x}) - H_{i,sq}^{\ell}(\mathbf{x})$$

- Criteria that rely only on the gradient marginal distributions of the GP.

Derivative-based Active Learning : existing work

Acquisition functions [1]

Maximum Variance :

$$\alpha_{\text{Var}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2$$

Variance Reduction :

$$\alpha_{\text{VarRed}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2 - \sigma^{\ell}_{i,sq}(\mathbf{x})^2$$

Entropy :

$$\alpha_{\text{Ent}}(\mathbf{x}) = \sum_{i=1}^{d} H_{i,sq}(\mathbf{x}) - H^{\ell}_{i,sq}(\mathbf{x})$$

- Criteria that rely only on the gradient marginal distributions of the GP.
- Local criteria : influence pointwise.

Acquisition functions [1]

Maximum Variance :

$$\alpha_{\text{Var}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2$$

Variance Reduction :

$$\alpha_{\text{VarRed}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2 - \sigma_{i,sq}^{\ell}(\mathbf{x})^2$$

Entropy :

$$\alpha_{\text{Ent}}(\mathbf{x}) = \sum_{i=1}^{d} H_{i,sq}(\mathbf{x}) - H_{i,sq}^{\ell}(\mathbf{x})$$

- Criteria that rely only on the gradient marginal distributions of the GP.
- Local criteria : influence pointwise.
- Uses a proxy for look-ahead variance.

Derivative-based Active Learning : existing work

Acquisition functions [1]

Maximum Variance :

$$\alpha_{\text{Var}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2$$

Variance Reduction :

$$\alpha_{\text{VarRed}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2 - \sigma^{\ell}_{i,sq}(\mathbf{x})^2$$

Entropy :

$$\alpha_{\text{Ent}}(\mathbf{x}) = \sum_{i=1}^{d} H_{i,sq}(\mathbf{x}) - H^{\ell}_{i,sq}(\mathbf{x})$$

- Criteria that rely only on the gradient marginal distributions of the GP.
- Local criteria : influence pointwise.
- Uses a proxy for look-ahead variance.
- No integration over the input pdf.

Derivative-based Active Learning : existing work

Acquisition functions [1]

Maximum Variance :

$$\alpha_{\text{Var}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2$$

Variance Reduction :

$$\alpha_{\text{VarRed}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2 - \sigma_{i,sq}^{\ell}(\mathbf{x})^2$$

Entropy :

$$\alpha_{\text{Ent}}(\mathbf{x}) = \sum_{i=1}^{d} H_{i,sq}(\mathbf{x}) - H_{i,sq}^{\ell}(\mathbf{x})$$

- Criteria that rely only on the gradient marginal distributions of the GP.
- Local criteria : influence pointwise.
- Uses a proxy for look-ahead variance.
- No integration over the input pdf.
- Don't apply to multidimensional
- and functional output.

Derivative-based Active Learning : existing work

Acquisition functions [1]

Maximum Variance :

$$\alpha_{\mathrm{Var}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2$$

Variance Reduction :

$$\alpha_{\mathrm{VarRed}}(\mathbf{x}) = \sum_{i=1}^{d} \sigma_{i,sq}(\mathbf{x})^2 - \sigma_{i,sq}^{\ell}(\mathbf{x})^2$$

Entropy :

$$\alpha_{\mathrm{Ent}}(\mathbf{x}) = \sum_{i=1}^{d} H_{i,sq}(\mathbf{x}) - H_{i,sq}^{\ell}(\mathbf{x})$$

- Criteria that rely only on the gradient marginal distributions of the GP.
- Local criteria : influence pointwise.
- Uses a proxy for look-ahead variance.
- No integration over the input pdf.
- Don't apply to multidimensional
- and functional output.

Derivative-based Active Learning : existing work

- Marginal-distribution-based criteria.
- Local criteria : pointwise influence.
- Uses a proxy for look-ahead variance.
- No integration over the input pdf.
- Not adapted to the multi-output case
- nor to functional output.

Derivative-based Active Learning : existing work

Gradient-based Active Learning for single output

- Marginal-distribution-based criteria.
- Local criteria : pointwise influence.
- Uses a proxy for look-ahead variance.
- No integration over the input pdf.
- Not adapted to the multi-output case
- nor to functional output.

Derivative-based Active Learning : existing work

Gradient-based Active Learning for single output

- Marginal-distribution-based criteria.
- Local criteria : pointwise influence.
- Uses a proxy for look-ahead variance.
- No integration over the input pdf.

Extension to multi-output

- Not adapted to the multi-output case
- nor to functional output.

Derivative-based Active Learning : existing work

Gradient-based Active Learning for single output

- Marginal-distribution-based criteria.
- Local criteria : pointwise influence.
- Uses a proxy for look-ahead variance.
- No integration over the input pdf.

Extension to multi-output

- Not adapted to the multi-output case

Functional outputs & Application to CATHY

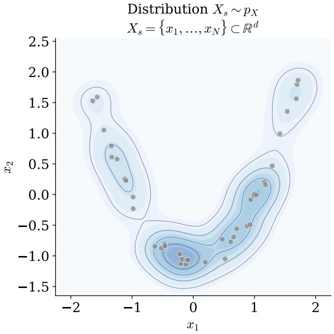
- nor to functional output.

Table of Contents

1. Gradient-based Active Learning for single output
2. Extension to multi-output.
3. Special case of functional output and application to CATHY
4. Conclusion

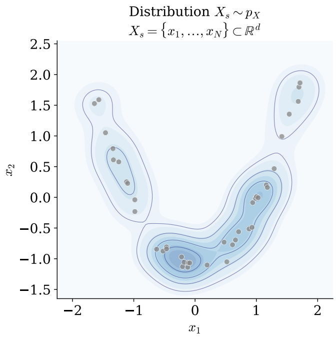
Gradient-based Active Learning

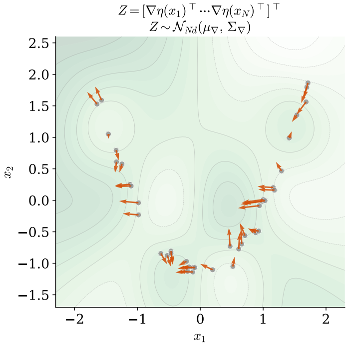
Given points $\mathbf{X_s} = (x_1, \ldots, x_N) \in \left(\mathbb{R}^d\right)^N$,



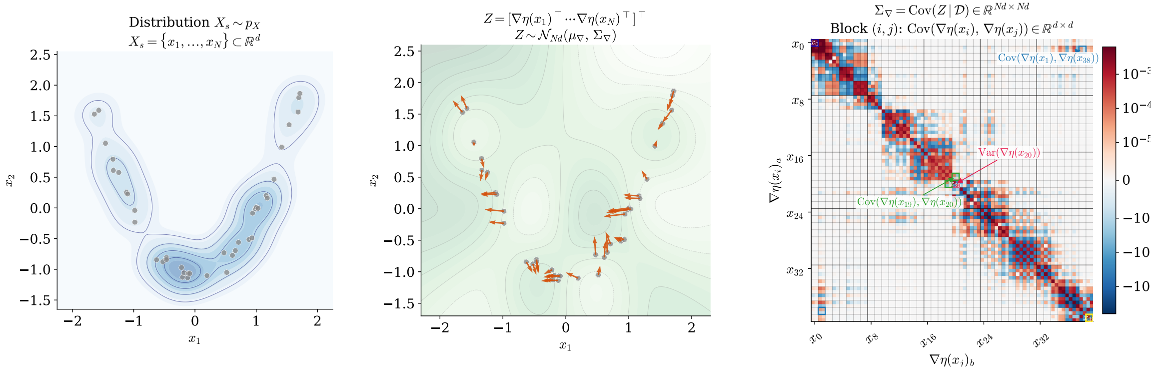
Gradient-based Active Learning

Given points $\mathbf{X_s} = (x_1, \ldots, x_N) \in \left(\mathbb{R}^d\right)^N$, the r.v. $Z := \nabla\eta(\mathbf{X_s})$





Given points $\mathbf{X_s} = (x_1, \ldots, x_N) \in \left(\mathbb{R}^d\right)^N$, the r.v. $Z := \nabla\eta(\mathbf{X_s}) \mid \mathcal{D} \sim \mathcal{N}_{Nd}(\mu_\nabla, \Sigma_\nabla)$.



- **Joint posterior:** $Z = [\nabla\eta(x_1)^\top \cdots \nabla\eta(x_N)^\top]^\top \sim \mathcal{N}_{Nd}(\mu_\nabla, \Sigma_\nabla)$ captures correlations across *points* and *coordinates* simultaneously.

Gradient-based Active Learning

Given points $\mathbf{X_s} = (x_1, \dots, x_N) \in \left(\mathbb{R}^d\right)^N$, the r.v. $Z := \nabla\eta(\mathbf{X_s}) \mid \mathcal{D} \sim \mathcal{N}_{Nd}(\mu_\nabla, \Sigma_\nabla)$.

Global one-step-ahead Variance Reduction (GlobalGradVarRed) [5]

Our aim is to find the point $\mathbf{x}^*$ that will reduce the most the one-step-ahead variance of $\|Z\|_2^2 = Z^\top Z$, i.e :

$$\alpha(\mathbf{x}^*) = \mathrm{Var}(Z^\top Z) - \mathbb{E}_y\left[\mathrm{Var}\left(Z^\top Z \mid \mathcal{D}, (\mathbf{x}^*, y)\right)\right]$$

with,

$$\mathrm{Var}(Z^\top Z) = 2\mathrm{Tr}(\Sigma_\nabla^2) + 4\mu_\nabla^\top \Sigma_\nabla \mu_\nabla$$

Gradient-based Active Learning

Given points $\mathbf{X_s} = (x_1, \ldots, x_N) \in \left(\mathbb{R}^d\right)^N$, the r.v. $Z := \nabla\eta(\mathbf{X_s}) \mid \mathcal{D} \sim \mathcal{N}_{Nd}(\mu_\nabla, \Sigma_\nabla)$.

Global one-step-ahead Variance Reduction (GlobalGradVarRed) [5]

Our aim is to find the point $\mathbf{x}^*$ that will reduce the most the one-step-ahead variance of $\|Z\|_2^2 = Z^\top Z$, i.e :

$$\alpha(\mathbf{x}^*) = \mathrm{Var}(Z^\top Z) - \mathbb{E}_y\left[\mathrm{Var}\left(Z^\top Z \mid \mathcal{D}, (\mathbf{x}^*, y)\right)\right]$$

with,

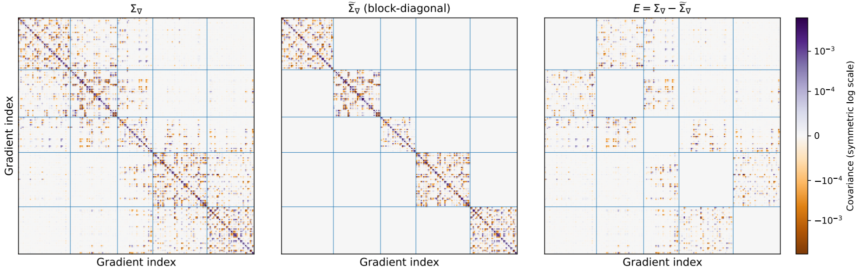
$$\mathrm{Var}(Z^\top Z) = 2\mathrm{Tr}(\Sigma_\nabla^2) + 4\mu_\nabla^\top \Sigma_\nabla \mu_\nabla$$

Time complexity	Memory complexity
$\mathcal{O}\left(N_f B\left(n^3 + Nn^2 d + N^2 n d^2\right)\right)$	$\mathcal{O}\left(N_f B(N^2 d^2 + Nn^2 + Nnd)\right)$

Scalability to high-dimensional input space

$\mathbf{X_s}$ is partitioned into C clusters $\{C_i\}_{i=1}^{C}$. Σ_∇ is approximated as $\tilde{\Sigma}_\nabla = \operatorname{diag}\left(\Sigma_\nabla^{(11)}, \ldots, \Sigma_\nabla^{(CC)}\right)$.

$$\widetilde{\mathrm{Var}}\left(Z^\top Z\right) = \sum_{i=1}^{C} \left(2\mathrm{Tr}\left(\Sigma_\nabla^{(ii)}\Sigma_\nabla^{(ii)}\right) + 4\mu_\nabla^{(i)\top}\Sigma_\nabla^{(ii)}\mu_\nabla^{(i)}\right)$$



Scalability to high-dimensional input space [5]

Bound on variance approximation

Let $\boldsymbol{E} = \Sigma_\nabla - \widetilde{\Sigma}_\nabla$ the error matrix. Under mild assumptions :

$$|V - \tilde{V}| \leq 2\|\boldsymbol{E}\|_F^2 + 4\|\mu_\nabla\|_2^2\|\boldsymbol{E}\|_2$$

Define: $\Delta_{ij} = \min_{x \in C_i, x' \in C_j} r(\mathbf{x}, \mathbf{x}')$, $\Delta = \min_{i \neq j} \Delta_{ij}$, and there exists $B > 0$, $h : \mathbb{R}^+ \to \mathbb{R}$ non-increasing such that :

$$\|\boldsymbol{E}\|_F^2 \leq \Big(N^2 - \sum_{i=1}^{C} n_i^2\Big) h(\Delta),$$

$$\|\boldsymbol{E}\|_2 \leq B\sqrt{h(\Delta)},$$

Numerical application on test functions

- Initial DoE : $n_0 = 5d$ with Sobol sequence
- Total budget : $n = 10d$ with 30 repetitions per active learning step
- Kernel function : ARD-Matérn 5/2
- Metrics : RMSE on DGSM estimation.

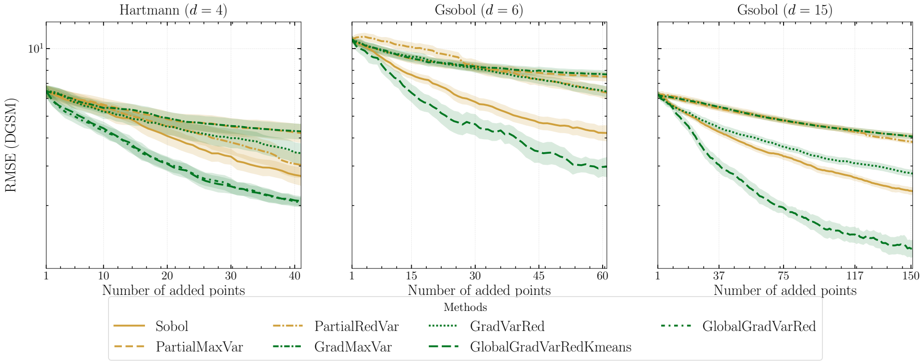


Table of Contents

1. Gradient-based Active Learning for single output
2. Extension to multi-output.
3. Special case of functional output and application to CATHY
4. Conclusion

Problem formulation

$$x_1 \sim \mathcal{U}(0,1) \quad x_1 \sim \mathcal{N}(\mu, \sigma^2) \quad \vdots \quad x_d \sim \beta(2,2)$$

$$\mathcal{M} : \mathbb{R}^d \to \mathbb{R}^p$$

$$y_1 = \mathcal{M}_1(\mathbf{x}) \quad y_2 = \mathcal{M}_2(\mathbf{x}) \quad \vdots \quad y_p = \mathcal{M}_p(\mathbf{x})$$

$\mathbf{x} \in \mathbb{R}^d$ $\qquad$ $\mathbf{y} \in \mathbb{R}^p$

Why is investigating multi-output interesting for CATHY?

CATHY involves several interacting processes:

- Water
- Surface / subsurface
- Erosion and pesticide

Gradient-based active learning for multi-output

For outputs in $\mathbb{R}^p$, define: $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_p)^\top : \mathbb{R}^d \to \mathbb{R}^p$

- Each $\eta_j : \mathbb{R}^d \to \mathbb{R}$ is an independent GP, $j = 1, \ldots, p$
- All η_j share the same mean and covariance structure as η
- Independence across outputs: $\eta_j \perp \eta_k$ for $j \neq k$

Global Variance Reduction for Multi-Output (*Jacques* 🍏)

Let $x \in \mathbb{R}^d$. Recall the Jacobian matrix :

$$\mathrm{Jac}_\eta(x) = (\partial_i \, \eta_j(x))_{1 \leq i \leq d, 1 \leq j \leq p}$$

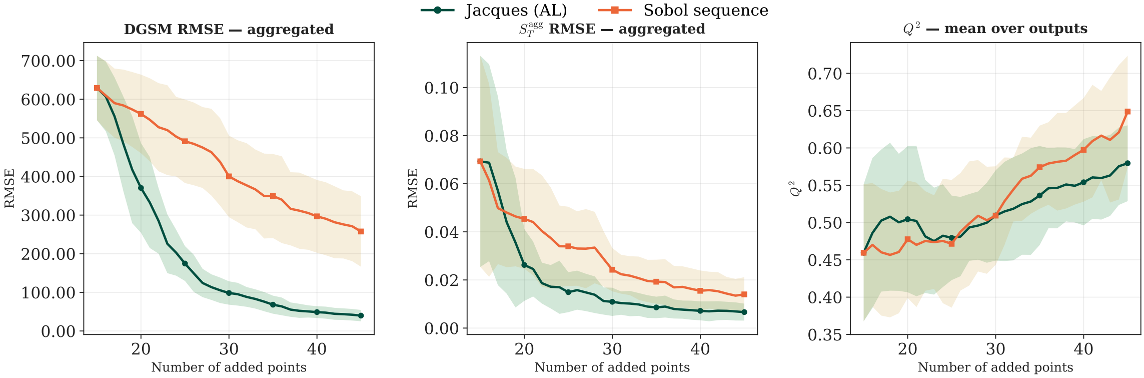
The Frobenius norm : $\|A\|_F^2 = \mathrm{tr}\left(A^\top A\right)$. Then, for a candidate $\mathbf{x}^* \in \mathbb{R}^d$:

$$\alpha(\mathbf{x}^*) = \mathrm{Var}\Big(\|Z\|_F^2\Big) - \mathbb{E}_y\left[\mathrm{Var}\Big(\|Z\|_F^2 \mid \mathcal{D} \cup \{(\mathbf{x}^*, y)\}\Big)\right]$$

with $Z = \mathrm{vect}\left(\mathrm{Jac}_\eta(\mathbf{X_s}) \mid \mathcal{D}\right) \in \mathbb{R}^{pNd} \sim \mathcal{N}_{pNd}(\mu, \Sigma)$.

Numerical experiment

- Test function : $\mathcal{M} : \mathbf{x} \in \mathbb{R}^3 \to (\text{Ishigami}(\mathbf{x}), \text{Branin}(\mathbf{x})) \in \mathbb{R}^2$

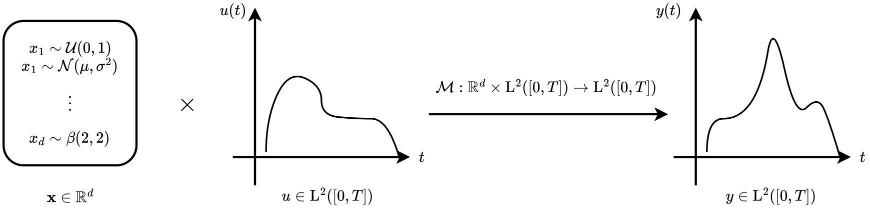


- Reduced uncertainty for estimating **DGSM** and **Sobol indices** using **Jacques**.
- Comparable mean prediction accuracy between rQMC and **Jacques**.

Table of Contents

1. Gradient-based Active Learning for single output
2. Extension to multi-output.
3. Special case of functional output and application to CATHY
4. Conclusion

How to tackle functional data ?

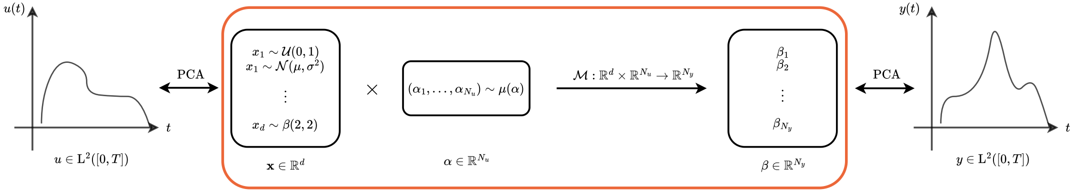


- Decompose functional data into a low-rank approximation using e.g PCA :

$$u(t) \approx \bar{u}(t) + \sum_{i=1}^{N_u} \alpha_i \, \varphi_i(t) \qquad y(t) \approx \bar{y}(t) + \sum_{j=1}^{N_y} \beta_j \, \phi_j(t)$$

- PCA scores are uncorrelated.

How to tackle functional data ?



- Decompose functional data into a low-rank approximation using e.g PCA :

$$u(t) \approx \bar{u}(t) + \sum_{i=1}^{N_u} \alpha_i \, \varphi_i(t) \qquad y(t) \approx \bar{y}(t) + \sum_{j=1}^{N_y} \beta_j \, \phi_j(t)$$

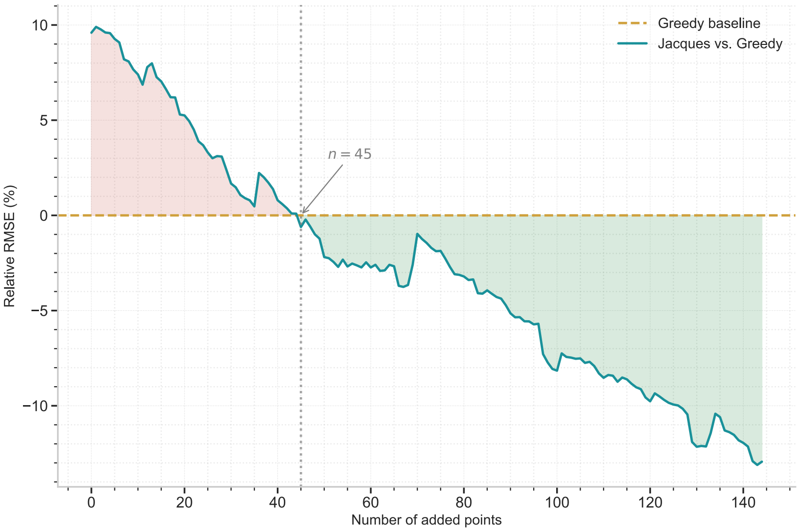
- PCA scores are uncorrelated.

It returns as a *simple* Vector to Multi-output problem !

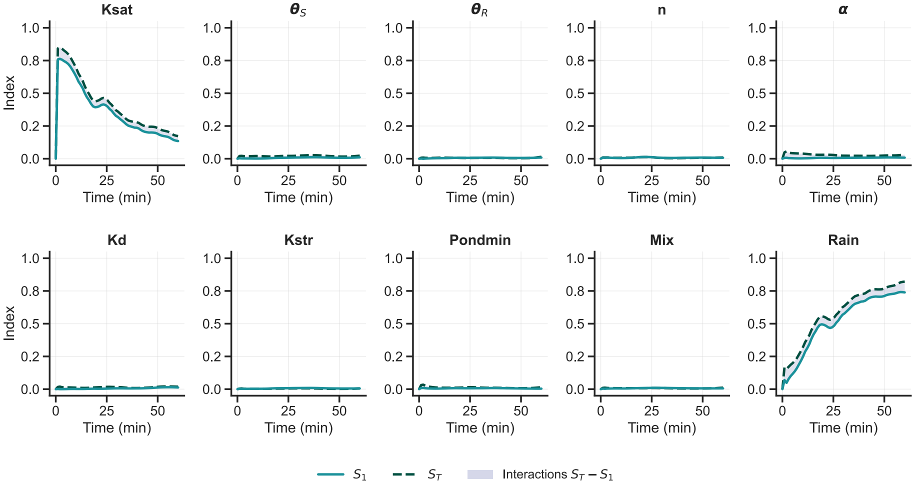
Application to CATHY

- Apply **Jacques** acquisition strategy to CATHY.
- PCA : $N_u = 3$ for Rain and $N_y = 7$ for Pesticide QoI (99% explained variance).
- Multi-output GP with Matérn 5/2
- Prediction quality : Relative RMSE with a full-budget upfront baseline model.
- Sensitivity metrics : Sobol indices for multidimensional and functional output [4].

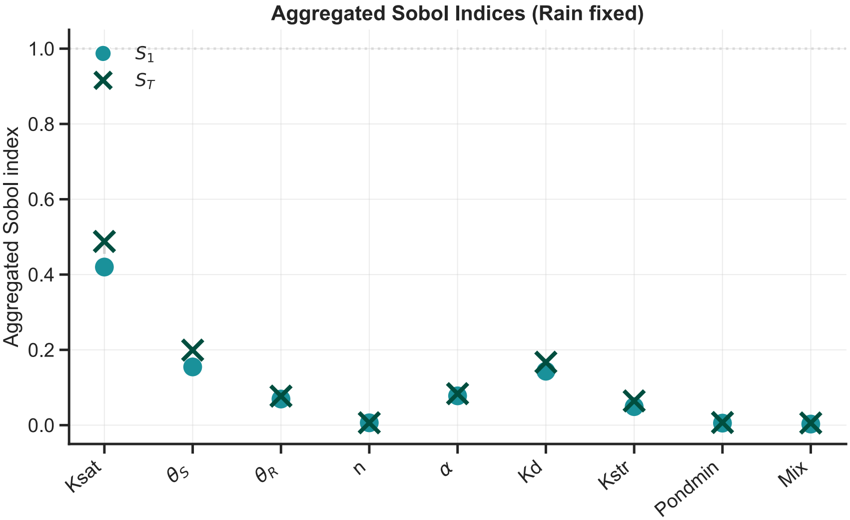
Numerical experiment



Sensitivity analysis



Sensitivity analysis at fixed rain



Sensitivity analysis at fixed rain

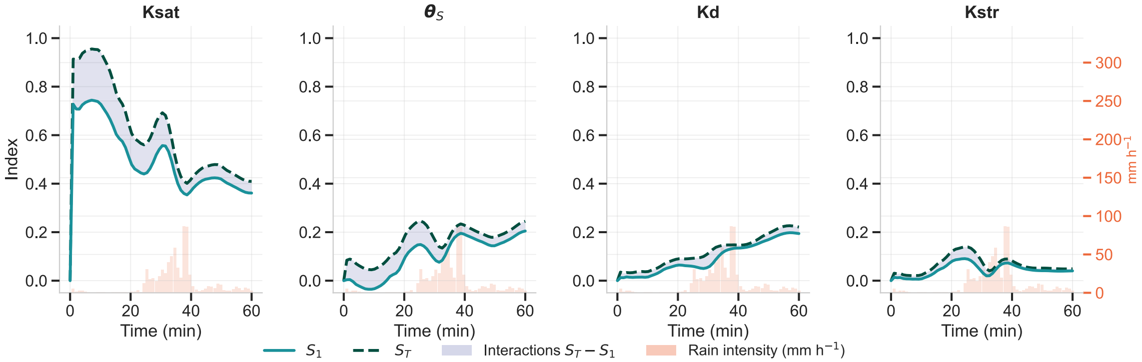


Table of Contents

1. Gradient-based Active Learning for single output
2. Extension to multi-output.
3. Special case of functional output and application to CATHY
4. Conclusion

Take-Aways

Key Points

1. Active Learning framework based on gradient in a context of complex input / output without gradient observation.
2. Application to a complex environment model and understanding pollutants transport.
3. Sensitivity analysis of CATHY integrating rainfall as input.
4. All developed methods relies on `torch` and benefit tensorization / auto-differentiation for time-efficient computations. The code for the proposed framework is available [1].

Ongoing work and open problems

- Adaptation of `Jacques` acquisition function and extension to the Vecchia-based approximation.
- Adaptation to Entropy-based acquisition.
- More test cases for multi-output and functional data with more methods.

[1] https://github.com/gulambert/RT_UQ_26

Thank you for your attention!

Questions & Discussion

[1] Thanks to the CATHY-Hotline from Quebec : Claudio Paniconi and Laura Gatel.

References I

[1] S. Belakaria et al. *Active Learning for Derivative-Based Global Sensitivity Analysis with Gaussian Processes*. 2024.

[2] M. Camporese et al. "Surface-subsurface flow modeling with path-based runoff routing, boundary condition-based coupling, and assimilation of multisource observation data". In: *Water Resources Research* 46.2 (2010). DOI: 10.1029/2008WR007536.

[3] S. Da Veiga et al. *Basics and Trends in Sensitivity Analysis*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2021. DOI: 10.1137/1.9781611976694.

[4] F. Gamboa et al. "Sensitivity analysis for multidimensional and functional outputs". In: *Electronic Journal of Statistics* 8.1 (2014), pp. 575–603. DOI: 10.1214/14-EJS895.

[5] G. Lambert, C. Helbert, and C. Lauvernet. *Gradient-based Active Learning with Gaussian Processes for Global Sensitivity Analysis*. 2026.

References II

[6] M. de Lozzo and A. Marrel. *Estimation of the derivative-based global sensitivity measures using a Gaussian process metamodel*. 2015.

[7] O. Roustant, F. Barthe, and B. Iooss. "Poincaré inequalities on intervals – application to sensitivity analysis". In: *Electronic Journal of Statistics* 11.2 (2017), pp. 3081–3119. DOI: 10.1214/17-EJS1310.

[8] I.M. Sobol and S. Kucherenko. "Derivative based global sensitivity measures". In: *Procedia - Social and Behavioral Sciences* 2.6 (2010). Sixth International Conference on Sensitivity Analysis of Model Output, pp. 7745–7746. ISSN: 1877-0428. DOI: 10.1016/j.sbspro.2010.05.208.

[9] S. Weill et al. "Coupling water flow and solute transport into a physically-based surface–subsurface hydrological model". In: *Advances in Water Resources* 34.1 (2011), pp. 128–136.