# Analyse de sensibilité globale pour des modèles numériques à sorties fonctionnelles.

Stage recherche de niveau M2

## 1 Informations pratiques

- Encadrants et contacts : Raphaël CARPINTERO PEREZ (raphael.carpinteroperez@cea.fr).
- Lieu : CEA - DAM Île-de-France, Bruyères-le-Châtel - 91297 Arpajon.
- Durée du stage : 5 à 6 mois avec financement.
- Début prévu du stage : début mars/avril 2027.
- Poursuite en thèse à envisager sous la direction de Sébastien DA VEIGA (sebastien.da-veiga@ensai.fr) et Guillaume PERRIN (guillaume.perrin@univ-eiffel.fr), en rattachement à l'école doctorale MATISSE à Rennes.

## 2 Contexte et motivation

La simulation numérique est un outil essentiel pour la conception ou la certification de systèmes complexes. De tels codes de calcul mettent en jeu de nombreux paramètres et sont souvent très coûteux. Afin d'obtenir une solution approchée à ces problèmes complexes, il est de plus en plus fréquent d'avoir recours à de la métamodélisation, qui consiste à créer des modèles de machine learning entraînés avec des bases données de simulations déjà réalisées. Dans ce contexte, la quantification des incertitudes est cruciale pour évaluer la fiabilité et la robustesse des prédictions, en prenant en compte les différentes sources d'incertitude. En particulier, l'analyse de sensibilité consiste à identifier quels paramètres ou variables exercent la plus grande influence sur les sorties du code ou du modèle, ces sorties étant généralement représentées comme des scalaires.

Il arrive cependant que ces sorties soient présentes sous des formes plus compliquées. Par exemple, pour chaque jeu de paramètres en entrée, on peut observer des séries temporelles ou des courbes. La figure 1 donne un exemple de jeu de données.

## 3 Objectifs du stage

L'objectif de ce stage est d'étudier différentes approches permettant de réaliser une analyse de sensibilité lorsque les sorties sont des courbes (par exemple des fonctions) plutôt que des variables scalaires.

1. **Indices ponctuels** : lorsque les sorties sont des séries temporelles, calculer un indice de sensibilité scalaire à chaque pas de temps $t_1, \cdots, t_D$.
2. **Réduction de dimension + agrégation** [6, 9] : réduire les sorties à quelques composantes latentes (ACP, ACP fonctionnelle...), calculer des indices de sensibilité sur ces composantes, puis les agréger pour revenir à un indice global.

3. **Indices de Sobol' Hilbertiens** [5] : définir directement la variance et les indices de Sobol' dans l'espace de Hilbert des sorties, généralisant le cadre scalaire classique sans réduction préalable.
4. **Approches à noyaux** [3, 4] : mesurer la dépendance entre entrées et sorties au moyen de noyaux définis directement dans l'espace des sorties.

Pour ce dernier point, se pose la question du choix du noyau. Deux courbes de formes proches mais ayant des décalages temporels sont jugées très éloignées en utilisant des distances $L^2$. On peut donc remplacer ces distances par des représentations spectrales (transformations de Fourier/ondelettes [8]), d'autres distances telles que des déformations temporelles dynamiques ou des noyaux avec alignement global [2], des distances élastiques partielles [1].

Un challenge du stage est de pousser plus loin le choix du noyau que l'on souhaite construire à partir de difféomorphismes entre courbes [7].

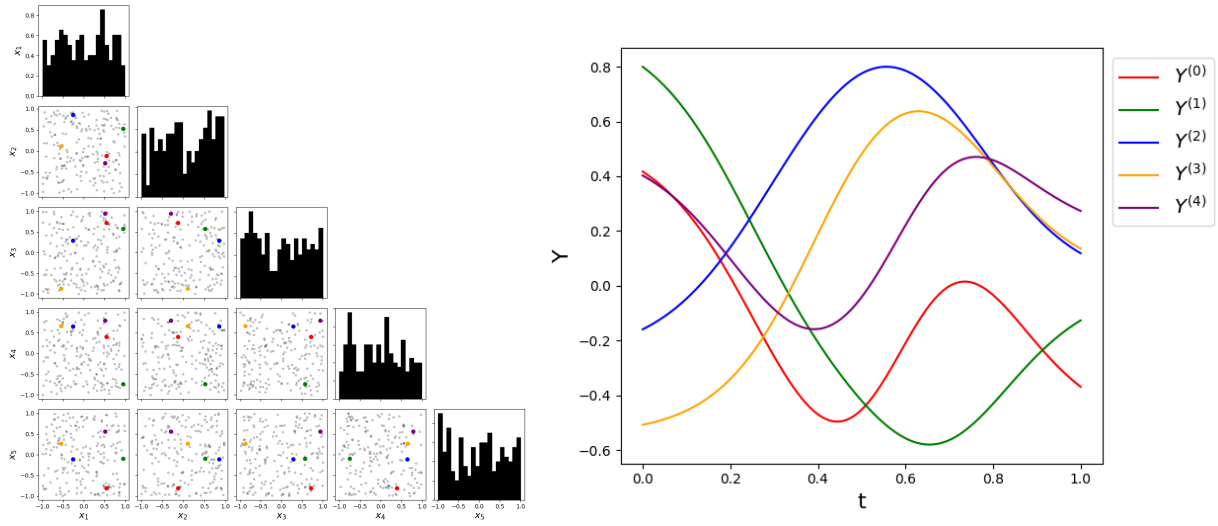


FIGURE 1 – Entrées vectorielles (gauche) et sorties fonctionnelles associcées (droite).

## Références


[1] Paul Castéras, Josselin Garnier, Julien Bect, and Gwenaël Salin. Bayesian calibration with functional outputs with elastic partial matching. 2026.

[2] Marco Cuturi. Fast global alignment kernels. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 929–936, 2011.

[3] Sebastien Da Veiga. Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation*, 85(7) :1283–1305, 2015.

[4] Sébastien Da Veiga. Kernel-based anova decomposition and shapley effects–application to global sensitivity analysis. *arXiv preprint arXiv :2101.05487*, 2021.

[5] Fabrice Gamboa, Alexandre Janon, Thierry Klein, and Agnès Lagnoux. Sensitivity analysis for multidimensional and functional outputs. 2014.

[6] Matieyendou Lamboni, David Makowski, Simon Lehuger, Benoit Gabrielle, and Hervé Monod. Multivariate global sensitivity analysis for dynamic crop models. *Field Crops Research*, 113(3) :312–320, 2009.

[7] Paul Lartaud and Gwenaël Salin. Diffeomorphic registration distances for bayesian calibration of infinite-dimensional computer models. *arXiv preprint arXiv :2605.10249*, 2026.

[8] Stéphane Mallat. *A wavelet tour of signal processing*, volume 2. Academic press San Diego, 1999.

[9] Amandine Marrel, Bertrand Iooss, Michel Jullien, Béatrice Laurent, and Elena Volkova. Global sensitivity analysis for models with spatially dependent outputs. *Environmetrics*, 22(3) :383–397, 2011.