

Analyse de l'échantillonnage préférentiel en grande dimension dans deux régimes d'évènement rare

Jason Beh^{1,2}, Florian Simatos², Jérôme Morio^{1,2}

¹ ONERA/DTIS, Université de Toulouse, F-31055 Toulouse

² Fédération ENAC ISAE-SUPAERO ONERA, Université de Toulouse, 31000 Toulouse

jason.beh@onera.fr

La durée prévue de thèse: janvier 2023 – décembre 2025

Abstract

Le problème d'estimation d'une probabilité faible est souvent abordé par échantillonnage préférentiel qui consiste à générer des échantillons selon une loi auxiliaire adaptée. Malgré son efficacité en petite dimension, l'échantillonnage préférentiel commet de fortes erreurs d'estimation lorsque la dimension de la variable d'entrée est grande, et devient très sensible au choix de la loi auxiliaire. Ce constat motive une étude théorique autour de l'échantillonnage préférentiel en grande dimension, un cadre peu analysé dans la littérature, afin de mieux comprendre les problématiques qui surviennent dans ce contexte.

D'un point de vue mathématique, l'étude en grande dimension que nous proposons consiste à faire tendre la dimension vers l'infini. Deux régimes d'évènement rare se distinguent alors: 1/ la probabilité à estimer reste bornée inférieurement hors de zéro, ou 2/ elle tend vers zéro avec la dimension. Le premier régime apparaît par exemple lorsque la probabilité à estimer fait intervenir un processus stochastique comme une variable d'entrée. Le processus stochastique peut être approché par une somme finie de variables aléatoires par analyse en composantes principales, et la probabilité issue devient également approchée. La dimension étant la taille de la somme, cette probabilité approchée converge vers la probabilité originale à estimer lorsque la dimension tend vers l'infini.

Le second régime où la probabilité à estimer tend vers zéro est le régime naturel à considérer mathématiquement lorsque l'on parle d'évènements rares. Cependant, ce régime n'est pas encore étudié de façon approfondie dans la littérature, et ce même en dimension finie, car le comportement des estimateurs peut être modifié de façon significative selon la vitesse de décroissance de la probabilité. Leur comportement se complexifie davantage si ce régime est couplé avec la dimension qui tend vers l'infini.

Nos avancées sur l'étude de l'échantillonnage préférentiel en grande dimension dans ces deux régimes seront présentées. À l'aide des théorèmes de convergence de suite triangulaire de variables aléatoires, nous formulons les conditions nécessaires et suffisantes pour la consistance d'estimateur d'échantillonnage préférentiel. Nous traduisons ces conditions vers la vitesse de croissance minimale de la taille d'échantillon avec la dimension dans des cas explicites en supposant que la loi initiale est la loi normale standard, et en considérant certaines lois auxiliaires normales. Nous discutons également les conditions pour que les estimateurs admettent un théorème central limite vers une loi limite gaussienne, et éventuellement les conditions pour d'autres loi limites.

Learning signals defined on graphs with optimal transport and Gaussian process regression

Raphaël Carpintero Perez^{†,1,2}, Sébastien Da Veiga^{§,3}, Josselin Garnier^{§,2}, Brian Staber^{§,1}

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Nov. 2022 – Nov. 2025

¹ Safran Tech, Digital Sciences & Technologies Department
{`raphael.carpintero-perez`, `brian.staber`}@safrangroup.com

² CMAP, Ecole polytechnique, Institut Polytechnique de Paris
{`raphael.carpintero-perez`, `josselin.garnier`}@polytechnique.edu

³ ENSAI, CREST
`sebastien.da-veiga@ensai.fr`

Abstract

Many problems in computational physics rely on solving partial differential equations (PDEs) using the finite element method involving meshes. Although much appreciated, they often require very fine discretizations of the meshes, which quickly become intractable. Meta-models are increasingly used to replace these costly methods as they allow to explore new geometries more easily. In this context, the supervised learning task involves inputs being graphs with continuous node attributes, different number of nodes and edges, and outputs being physical quantities of interest defined on each node.

While some methods like Graph Neural Networks are intrinsically designed to predict signals defined on graphs or point clouds, a natural question is the extension of general scalar output regression models to such complex outputs. Changes between input geometries in terms of both size and adjacency structure in particular make this transition non-trivial.

In this talk, we present an innovative strategy for Gaussian process regression where inputs are large graphs with continuous node attributes and outputs are signals defined on the nodes of the associated inputs. The methodology relies on the combination of regularized optimal transport, dimension reduction techniques, and the use of Gaussian processes indexed by graphs [1]. In addition to enabling signal prediction, the main point of our proposal is to come with confidence intervals on node values, which is crucial for uncertainty quantification. Numerical experiments highlight the efficiency of the method to solve real problems in fluid dynamics and solid mechanics.

References

- [1] Raphaël Carpintero Perez, Sébastien Da Veiga, Josselin Garnier, and Brian Staber. Gaussian process regression with sliced Wasserstein Weisfeiler-Lehman graph kernels. In *Proceedings of The 27th International Conference on Artificial Intelligence and Statistics*, pages 1297–1305, 2024.

Krigeage adaptatif de modèles numériques séquentiels pour l'étude de la fiabilité de systèmes sollicités en fatigue

Contexte

Le travail présenté dans cette communication se place dans le contexte de l'étude de la durabilité de systèmes, dont le résultat correspond à un cumul d'endommagement, noté $D \in [0,1]$. Cette grandeur d'intérêt est la sortie d'une chaîne numérique composée séquentiellement d'un premier modèle noté m_A et d'un second modèle noté m_B . Dans le cas des systèmes industriels complexes, le modèle m_A correspond à la résolution du problème dynamique par éléments finis du système complet. Le temps de calcul associé est généralement sévère (jusqu'à 30 heures pour une seule exécution) et la sortie $m_A(\mathbf{X}_A)$ de ce calcul est complexe : il s'agit d'un tenseur des contraintes dépendant du temps et défini en chaque nœud du maillage. Le modèle m_B permet ensuite de cartographier l'endommagement dans les zones dites à risque d'initiation de fissure à partir du post-traitement des résultats du modèle m_A . Son évaluation dure environ 3 minutes pour une seule localisation, cas que l'on considèrera dans cette présentation. Les vecteurs $\mathbf{X}_A \in \mathbb{R}^{n_A}$ et $\mathbf{X}_B \in \mathbb{R}^{n_B}$ représentent les entrées de m_A et m_B respectivement et sont considérés comme aléatoires. Notons que \mathbf{X}_A n'est l'entrée que de m_A , et \mathbf{X}_B n'intervient qu'au niveau de m_B , la chaîne a donc une propriété de séparation des variables d'entrée, que la figure 1 permet de visualiser.

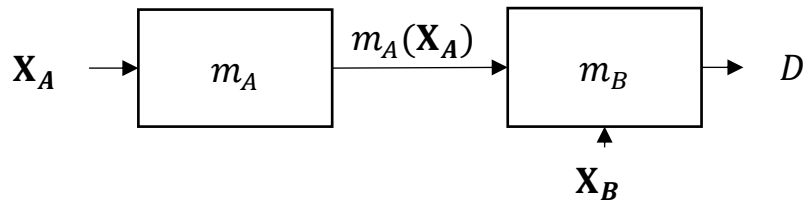


Figure 1 : Principe du chaînage séquentiel des modèles

L'objectif est alors d'estimer la probabilité que le dommage D , évalué dans une zone à risque, dépasse une valeur seuil d_{seuil} définie par le cahier des charges. Elle est définie telle que :

$$p_f = P(d_{seuil} - D(\mathbf{X}_A, \mathbf{X}_B) \leq 0) \quad (1)$$

Avec $D(\mathbf{X}_A, \mathbf{X}_B) = m_B \circ m_A(\mathbf{X}_A, \mathbf{X}_B) = m_B(m_A(\mathbf{X}_A), \mathbf{X}_B)$.

La fonction de performance associée à la probabilité de défaillance s'écrit ainsi comme :

$$g(\mathbf{X}_A, \mathbf{X}_B) = d_{seuil} - m_B(m_A(\mathbf{X}_A), \mathbf{X}_B) \quad (2)$$

L'estimation de p_f par l'utilisation directe des méthodes de simulation de type Monte Carlo, n'est ici pas possible compte-tenu du nombre d'appels à g qu'elles provoquent. Dans des configurations similaires, des approches non intrusives basées sur la substitution du modèle coûteux g par un métamodèle \hat{g} ont fait leurs preuves (Maliki Moustapha, 2022). On peut par exemple citer les méthodes qui combinent krigeage adaptatif et simulations pour l'estimation de probabilités de défaillance (AK-MCsi) (Lelièvre, 2018).

Travaux développés et contenu de la présentation

Une application directe de la méthode AK-MCS pouvant engendrer un nombre d'appels à g trop important au regard du budget en temps de calcul fixé par l'industriel, une nouvelle adaptation nommée Active Kriging for Sequential Model (AK-SM) est proposée. AK-SM exploite deux caractéristiques propres à la chaîne numérique :

1. Le coût c_A de m_A est 600 à 1000 fois plus élevé que le coût c_B de m_B ;
2. La séparation des variables d'entrée de la chaîne séquentielle de modèles.

La méthode proposée s'appuie sur la séparation des variables et sur une décomposition fonctionnelle tronquée $g(\mathbf{X}_A, \mathbf{X}_B) = g_0 + g_A(\mathbf{X}_A) + g_B(\mathbf{X}_B)$, dont les composantes g_A et g_B sont évaluées par projection de \mathbf{X} sur deux hyperplans passant par un point d'ancrage $\bar{\mathbf{X}} = (\bar{\mathbf{X}}_A, \bar{\mathbf{X}}_B)$. La décomposition fonctionnelle est ensuite utilisée pour réduire le nombre d'appels à m_A lors des itérations de la procédure AK-MCS en imputant parfois la quantité complexe et coûteuse $m_A(\mathbf{X}_A)$ par $m_A(\bar{\mathbf{X}}_A)$. Ainsi, certains calculs n'ont alors plus qu'un coût c_B , ce qui entraîne une diminution du coût total pour la calibration du métamodèle \hat{g} . L'application de AK-SM sur un cas analytique ainsi que sur un cas éléments finis a démontré un potentiel de réduction du coût de calculs d'un facteur 4 par rapport à AK-MCS, ce qui est significatif dans le contexte industriel de la thèse.

Des résultats seront présentés sur un cas analytique et sur un cas éléments finis simplifié mais représentatif du cas industriel. Le cas analytique sera un problème de fiabilité traitant du phénomène d'interaction entre la fatigue et le fluage, et le cas éléments finis « intermédiaire » traitera de la fiabilité en fatigue multiaxiale d'une équerre de fixation sollicitée par accélérations à la base.

Nom, encadrement et affiliations

Doctorant : Thomas CONSTANT¹ thomas.constant@knds.fr

Encadrement académique : Cécile MATTRAND², Nicolas GAYTON⁴

Encadrement entreprise : Laëtitia FOUCHE³

¹Université Clermont Auvergne/KNDS, Clermont Auvergne INP, CNRS, Institut Pascal, F-63000, Clermont-Ferrand, France

² Université Clermont Auvergne, Clermont Auvergne INP, CNRS, Institut Pascal, F-63000, Clermont-Ferrand, France

³ KNDS, 18000, Bourges, France

⁴ Université Clermont Auvergne, Clermont Auvergne INP, CNRS, Institut Pascal, F-63000, Clermont-Ferrand, France

Début de thèse : 01/10/2022

Fin de thèse : 10/2025

Références

Lelièvre, 2018. Développement des méthodes AK pour l'analyse de fiabilité: Focus sur les événements rares, s.l.: Université Clermont Auvergne: Génie mécanique [physics.class-ph]..

Maliki Moustapha, S. M. B. S., 2022. Active learning for structural reliability: Survey, general framework and benchmark. *Structural Safety*.

Sensitivity Analysis in Constrained Bayesian Optimization with Uncertainties

Noé Fellmann^{1,2}

Céline Helbert¹, Christophette Blanchet¹

Adrien Spagnol², Delphine Sinoquet²

¹Ecole Centrale de Lyon

²IFP Energies nouvelles

Thèse : 01/12/2021 - 30/11/2024

This study tackles the challenge of chance-constrained optimization under uncertainties, which entails significant computational burdens in practical applications. We utilize EFISUR [1], an existing algorithm for Bayesian optimization. It relies on modeling the problem using Gaussian process regression in the joint space and defining an acquisition criterion in the joint optimized-uncertain input space that considers both the average improvement in the objective function and the reliability of the constraints. However, high dimensionality in either the design space or the uncertain parameter space can pose challenges due to the complexity of the optimization steps and Gaussian Processes (GPs) modeling.

To address this issue, we propose different strategies to integrate sensitivity analysis (SA) into EFISUR. Through comprehensive numerical test cases, we enhance the efficacy of EFISUR and enable its applicability in higher dimensions.

Keywords— Robust optimisation under constraints, Bayesian optimization, Sensitivity analysis

Références

- [1] Reda El AMRI, Rodolphe Le RICHE, Céline HELBERT, Christophette BLANCHET-SCALLIET et Sébastien Da VEIGA. “A Sampling Criterion for Constrained Bayesian Optimization with Uncertainties”. In : *The SMAI Journal of computational mathematics* 9 (2023), p. 285-309. DOI : 10.5802/smai-jcm.102.

Bayesian calibration of a steam generator clogging simulation code

E. Jaber^{†,1,2,3}, E. Remy^{†,1}, V. Chabridon^{†,1}, M. Mougeot^{§,2}, D. Lucor^{§,3}

[†] PhD student (presenting author) [‡] PhD industrial advisors [§] PhD supervisors

PhD expected duration: Feb. 2023 – Feb. 2026

¹ EDF R&D, PRISME (Performance, Industrial Risk, Monitoring for Maintenance and Operation) department, 6 quai Watier, 78401 Chatou, France

² Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, 91190 Gif-sur-Yvette, France

³ Laboratoire Interdisciplinaire Des Sciences Du Numérique, LISN-CNRS, F-91403 Orsay, France

Abstract

Clogging of steam generators (SGs) in pressurized water reactors is a complex phenomenon that develops over extended operational periods. To maintain the asset's efficiency and overall performance, chemical cleaning is conducted during outages. To improve predictions for maintenance scheduling, a physical model and the computer code THYC-Puffer-DEPOTHYC [2] have been developed, simulating clogging kinetics over long operational periods. Previous work [3] has identified several uncertain input variables of this code, and sensitivity analysis was conducted to evaluate their impact over the simulation time. One such variable is the calibration parameter related to the vena contracta phenomenon. Field data on clogging for specific SGs are typically collected through televised camera inspections or indirect measurements using Foucault currents, though this data is often limited or of low quality. To enhance the available dataset, EDF R&D has developed statistical regression models, specifically the ESTICOL tool, based on operational data. We propose a hybrid approach to improve the predictive accuracy of the computer model by applying Bayesian calibration [4] to the vena contracta parameter across different operational scenarios, utilizing various types of clogging field data. This method enables a more precise estimation of the Remaining Useful Life [1] of the SG asset, aiding in prognostics and maintenance planning.

References

- [1] Luca Biggio and Iason Kastanis. Prognostics and Health Management of Industrial Assets: Current Progress and Road Ahead. *Frontiers in Artificial Intelligence*, 3:578613, November 2020.
- [2] Q. Feng, J. Nebes, M. Bachet, S. Pujet, D. You, and E. Deri. Tube support plates blockage of PWR steam generators: thermohydraulics and chemical modeling, September 2023.
- [3] E. Jaber, V. Chabridon, E. Remy, M. Baudin, D. Lucor, M. Mougeot, and B. Iooss. Sensitivity analyses of a multi-physics long-term clogging model for steam generators. *International Journal for Uncertainty Quantification*, 2024.
- [4] Marc C. Kennedy and Anthony O'Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, 2001.

On the Bayesian calibration of a two-phase flow model under uncertain inlet conditions

S. Janati Idrissi^{1,3}, P.M. Congedo¹, O. Le-Maître², M.G. Rodio³

PhD started on : October 2022

Email: sanae.janatiidrissi@cea.fr

¹ Inria, CMAP, Institut Polytechnique de Paris, France

² CNRS, CMAP, Institut Polytechnique de Paris, France

³ CEA, ISAS, Laboratoire de Calcul intensif et d'Analyse Numérique, France

Abstract

Reliable predictions of two-phase flows are crucial for the safety and efficiency analysis of nuclear reactors. However, two-phase flow simulations are often subject to inherent uncertainties that arise from different sources including measurement uncertainty, parameter uncertainty, and model discrepancy. In this work, we focus on model calibration to infer model parameters in order to improve the predictive accuracy of numerical codes. We rely on a Bayesian technique called the Full Maximum a posteriori approach where a statistical model error is identified for every possible model parameter value. We investigate the impacts of uncertain inlet conditions on model calibration and uncertainty analysis through a case study of upward bubbly flow in a square duct. Particularly, we are interested in calibrating the coefficients in the coalescence and breakup terms in the interfacial area transport equation of the Two-Phase Regime Transition (TRITON) model implemented in Neptune CFD when the inlet void fraction and injection velocity are uncertain.

Robustness Analysis for Uncertainty Quantification of Computer Codes Using the Fisher-Rao Distance

B. Ketema^{†,1,2}, F. Gamboa^{§,1}, F. Costantino^{§,1},
R. Sueur^{§,2}, N. Bousquet^{§,2,3,4}, B. Iooss^{§,1,2}

[†] PhD student, [§] PhD supervisors

PhD expected duration: Feb. 2023 – Feb. 2026

¹ Institut de Mathématiques de Toulouse, Paul Sabatier University,

² EDF R&D,

³ LPSM, Sorbonne University, ⁴ SINCLAIR AI Laboratory, EDF R&D

In this PhD thesis, we focus on a quantity of interest (QoI) defined on the output Y of a numerical model G when the inputs X_1, \dots, X_d are considered to be random variables with reference laws f_1, \dots, f_d . The goal is to study the robustness of this QoI on Y with respect to *perturbations* on the input distributions f_i . In this work, we define the perturbation method using the Fisher-Rao distance on parametric families $\mathcal{P}_i := \{f_{i,\theta}\}_{\theta \in \Theta_i}$ containing f_i , as proposed in [3]. This requires to compute the information geometry of \mathcal{P}_i , mainly the geodesics for the Fisher information metric. For this presentation, we will expose the two main research topics we focused on until now.

The first research topic consists in taking into account the truncated nature of some of the inputs in the distributional perturbation method. This is important as some inputs X_i of G can be modeled, for instance, by a normal distribution $\mathcal{N}(\mu_0, \sigma_0^2)$ truncated on an interval $[a, b]$. This means that f_i , the law of X_i , is defined as a truncated normal density function

$$f_i := \frac{P(\mu_0, \sigma_0)}{N(\mu_0, \sigma_0)} \mathbf{1}_{[a, b]},$$

where $N(\mu_0, \sigma_0) := \int_a^b p(\mu_0, \sigma_0)(x) dx$ is the normalizing constant and $p(\mu_0, \sigma_0)$ is the density function of $\mathcal{N}(\mu_0, \sigma_0^2)$. We therefore computed the information geometry of such truncated families (as well as non-truncated ones). These are for instance the normal, log-normal, Gumbel, triangular and a few more.

The goal for the second research topic is to build confidence intervals for the QoI on Y . Here, we only consider the case where the QoI is given by an α -quantile. The particularity of this work is twofold. First, we specifically want the confidence intervals to be non-asymptotic. Second, the estimation method for the output quantile $q_\alpha(Y)$ after perturbation is based on an importance sampling (IS) procedure. Therefore, standard methods such as Wilks method or central limit theorem do not apply. Inspired by some techniques in [2] and a concentration inequality in [1], we built a non-asymptotic confidence interval for the quantile $q_\alpha(Y)$. In practice, this approach leads to an overly conservative confidence level. Nevertheless, an improvement is expected assuming boundedness on the likelihood ratio from the IS procedure.

References

- [1] S. Chatterjee and P. Diaconis. The sample size required in importance sampling. *The Annals of Applied Probability*, 28(2):1099–1135, 2018.
- [2] M. Ducoffe, S. Gerchinovitz, and J. S. Gupta. A high probability safety guarantee for shifted neural network surrogates. In *SafeAI@ AAAI*, pages 74–82, 2020.
- [3] C. Gauchy, J. Stenger, R. Sueur, and B. Iooss. An information geometry approach to robustness analysis for the uncertainty quantification of computer codes. *Technometrics*, 64(1):80–91, 2022.

Quantization-based Latin hypercube sampling for dependent inputs with an application to sensitivity analysis of environmental models.

Guerlain Lambert^{†,1}, Céline Helbert^{§,1}, Claire Lauvernet^{§,2}

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Oct. 2023 – Sep. 2026

¹ Institut Camille Jordan, CNRS UMR 5208, École Centrale de Lyon, Écully, France
`{guerlain.lambert,celine.helbert}@ec-lyon.fr`

² INRAE, RiverLy, 69625 Villeurbanne, France
`claire.lauvernet@inrae.fr`

Abstract

Les modèles numériques sont essentiels pour comprendre les phénomènes physiques complexes dans différents domaines. Pour gérer leur complexité, l'analyse de sensibilité, en particulier le criblage, est cruciale pour identifier les paramètres d'entrée influents. Les méthodes basées sur les noyaux, telles que le critère d'indépendance de Hilbert Schmidt (HSIC) [1], sont très utiles pour analyser les dépendances entre les entrées et les sorties. La mise en œuvre du critère d'indépendance de Hilbert Schmidt requiert des données provenant du modèle original, ce qui nécessite des stratégies d'échantillonnage efficaces afin de limiter le nombre de simulations numériques coûteuses. Alors que, pour les variables d'entrée indépendantes, les méthodes d'échantillonnage existantes telles que le Latin Hypercube Sampling (LHS) [2] sont efficaces pour estimer l'HSIC avec une variance réduite, l'incorporation de la dépendance est un défi. Cette présentation décrit une nouvelle variante du LHS, l'échantillonnage Quantization-based LHS (QLHS) [3], qui s'appuie sur la quantification vectorielle de Voronoï pour prendre en compte les entrées dépendantes. Cette méthode permet de bien couvrir la gamme des variations des variables d'entrée.

Cette présentation examine les estimateurs d'espérance basés sur le QLHS dans divers contextes de dépendance, en démontrant leur absence de biais. La méthode est appliquée à plusieurs modèles de complexité croissante, d'abord sur des exemples simples pour illustrer la théorie, puis sur des modèles hydrologiques environnementaux plus complexes, lorsque la dépendance est connue ou non, et avec des processus et des facteurs de plus en plus interactifs. La dernière application porte sur le digital twin d'un bassin versant viticole français (région du Beaujolais) afin de concevoir une bande filtrante végétale et de réduire les transferts d'eau, de sédiments et de pesticides des champs vers la rivière. Le QLHS est utilisé pour calculer les mesures HSIC et les tests d'indépendance, démontrant ainsi son utilité, en particulier dans le contexte de modèles complexes.

References

- [1] A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf. “Measuring Statistical Dependence with Hilbert-Schmidt Norms”. In: vol. 3734. 2005. ISBN: 978-3-540-29242-5. DOI: 10.1007/11564089_7.
- [2] M. D. McKay, R. J. Beckman, and W. J. Conover. “A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code”. In: *Technometrics* 21.2 (1979). Publisher: [Taylor & Francis, Ltd., American Statistical Association, American Society for Quality], pp. 239–245. ISSN: 0040-1706. DOI: 10.2307/1268522.
- [3] Guerlain Lambert, Céline Helbert, and Claire Lauvernet. *Quantization-based LHS for dependent inputs : application to sensitivity analysis of environmental models*. 2024.

Mitigating Overconfidence and Cost in Bayesian Field Inference

N. Polette^{†,1,2}, P. Sochala^{§,1}, A. Gesret^{§,2}, O. Le Maître^{§,3}

[†] PhD student (presenting author) [§] PhD supervisor

PhD expected duration: Nov. 2022 – Oct. 2025

¹ CEA, DAM, DIF, F-91297 Arpajon, France
`{nadege.polette,pierre.sochala}@cea.fr`

² MINES Paris, PSL University, Centre de Géosciences, France
`alexandrine.gesret@minesparis.psl.eu`

³ Centre de Mathématiques appliquées, CNRS and INRIA, Ecole Polytechnique
`olivier.le-maitre@polytechnique.edu`

Abstract

In Bayesian field inference, the objective is to estimate a scalar field and its uncertainties based on indirect observations. In general, for a given field, the observations are obtained through a forward model that is costly to evaluate, such as a system of partial differential equations. The approximation of the field posterior distribution through Markov chain Monte-Carlo sampling requires the use of a surrogate model to replace the expensive forward model. In addition, the infinite dimensionality of the field necessitates to develop a dimension reduction in order to deal with a finite dimensional parameter space. In a previous study, we used Polynomial Chaos Expansions (PCEs) and derived a novel parametrization of the field to address these challenges. However, the PCEs were computed during an offline process on a prior training set. In this presentation, we illustrate the limitations of this construction and provide some literature solutions to improve the approximations. These aspects are presented on a two-dimensional toy example.

Joint work with P. Sochala, O. Le Maître, A. Gesret

Plan d'expériences optimal sous incertitudes de campagnes d'essais de ténacités

Anthony Quintin^{a,d}, Rudy Chocat^a, Tom Petit^b, Cécile Mattrand^c, Jean-Marc Bourinet^d

^aCEA-DES, Université Paris-Saclay, Gif-sur-Yvette, 91190, France

^bCEA-DAM, Centre de Gramat, Gramat, 46500, France

^cUCA, Clermont Auvergne INP, Institut Pascal, Clermont-Ferrand, 63000, France

^dUCA, Clermont Auvergne INP, LIMOS, Clermont-Ferrand, 63000, France

Mail: anthony.quintin@cea.fr

Date début-fin thèse: Octobre 2022 - Septembre 2025

Abstract

La fragilisation par irradiation de l'acier des cuves des réacteurs à eau pressurisée constitue une priorité majeure des programmes de surveillance. Au cours des dernières décennies, la méthodologie de la Master Curve, développée par Wallin [1], est devenue une approche mondialement acceptée dans l'industrie nucléaire pour évaluer directement la ténacité des aciers au cours de la vie du réacteur. Cette évaluation permet ainsi de quantifier le décalage de la température de transition ductile-fragile T_0 en fonction de l'irradiation. Désormais normalisée par la norme ASTM E1921 [2], la méthodologie Master Curve nécessite la réalisation d'essais de ténacité sur des éprouvettes pré-fissurées de type CT (compact tension) aux dimensions standardisées. Ces essais doivent être menés à une ou plusieurs températures, situées dans la plage de transition ductile-fragile. Les données obtenues lors d'une campagne d'essais sont ensuite analysées pour déterminer la température de référence T_0 .

Cependant, l'estimation précise de T_0 reste un défi majeur dans l'approche Master Curve. Ce paramètre, en théorie intrinsèque à chaque matériau, démontre en réalité une sensibilité significative à plusieurs paramètres. Parmi ceux-ci figurent la variabilité intrinsèque à la rupture fragile (fonction de la distribution volumique des défauts), au nombre d'essais (limité en pratique), aux incertitudes géométriques subies, ou encore les températures d'essai sélectionnées par l'expérimentateur. La variabilité de T_0 aux températures d'essais peut atteindre des écarts jusqu'à ± 10 °C [3] avec des éprouvettes standards. Le matériau irradié se révélant de plus en plus précieux, une solution d'économie envisagée est l'utilisation d'éprouvettes miniaturisées mini-CT. Cette miniaturisation exacerbe la dépendance de T_0 aux paramètres d'essais, et notamment aux températures d'essai montant jusqu'à des variations de ± 15 °C [4] en raison du nombre plus élevé d'essais censurés (ténacité à la rupture supérieure à une limite analytique).

La question du choix des températures d'essai optimales au cours de la campagne d'essai pour réduire l'incertitude sur T_0 est donc de la plus grande importance. Nous proposons alors une méthode permettant d'évaluer le plan d'expériences optimal, i.e. la séquence optimale de températures d'essai réduisant la variabilité de T_0 . Cette méthode repose sur un problème d'optimisation, visant à réduire des indicateurs de dispersion de T_0 avec par exemple des fonctions d'utilités [5]. Les essais de ténacité sont reproduits par un modèle numérique d'éléments finis basé sur l'approche locale de la rupture fragile de Beremin. Les campagnes d'essais sont virtuellement générées par l'approche numérique proposée dans [6] pour évaluer T_0 par approche bayésienne. Les résultats des plans d'expériences optimaux pour un acier identifié semblent prometteurs et robustes selon la géométrie d'éprouvette utilisée. À terme, l'objectif serait d'étendre la démarche aux campagnes de résilience sur éprouvettes Charpy.

Keywords: Optimisation, Plan d'expériences, Fonctions d'utilité, Ténacité, Master Curve

References

- [1] K. Wallin, The scatter in KIC-results, Engineering Fracture Mechanics 19 (6) (1984) 1085–1093.
- [2] ASTM E1921-22a, ASTM E1921-22a, Standard Test Method for Determination of Reference Temperature T_0 for Ferritic Steels in the Transition Range (2022) West Conshohocken, PA, USA doi : 10.1520/E1921-22A.
- [3] K. Wallin, Master curve analysis of ductile to brittle transition region fracture toughness round robin data, The "EURO" fracture toughness curve, VTT manufacturing Technology (1998) 1455–0849.
- [4] M. Sánchez, S. Cicero, M. Kirk, E. Altstadt, W. Server, M. Yamamoto, Using mini-ct specimens for the fracture characterization of ferritic steels within the ductile to brittle transition range: A review, Metals 13 (1) (2023). doi : 10.3390/met13010176.
- [5] K. Chaloner, I. Verdinelli, Bayesian experimental design: A review, Statistical science (1995) 273–304.
- [6] A. Quintin, T. Petit, R. Chocat, C. Mattrand, J.-M. Bourinet, Uncertainty quantification of the reference temperature of reactor pressure vessel steel with experimental and numerical computation of fracture toughness tests, in: CFRAC 2023-The Seventh International Conference on Computational Modeling of Fracture and Failure of Materials and Structures, 2023.

Analyse de sensibilité pour systèmes multidisciplinaires

Edimah Songo – IFPEN/ONERA/ENSAI

1^{ère} année – 2 Novembre 2023 - 31 Octobre 2026

edimah.songo@ifpen.fr

Directeur de thèse Sébastien Da Veiga (ENSAI)

Encadrants Reda El Amri, Adrien Spagnol (IFPEN), Mathieu Balesdent, Loïc Brevault (ONERA)

Mots clés : Analyse de sensibilité, systèmes multidisciplinaires, conception de systèmes complexes, métamodélisation

La modélisation de systèmes complexes *multidisciplinaires*, comme les véhicules aérospatiaux et les moteurs électriques, présente des défis, notamment en matière d'optimisation et d'évaluation de performance. Ces systèmes impliquent des interactions entre de multiples disciplines parfois couplées, augmentant leur complexité. Pour adresser ces enjeux, cette thèse se concentre sur le développement de méthodes d'analyse de sensibilité adaptées aux systèmes multidisciplinaires. L'objectif est de déterminer les variables d'optimisation et les couplages interdisciplinaires les plus influents pour réduire l'espace de recherche et simplifier le problème d'optimisation, tout en préservant des performances optimales.

La recherche a débuté par l'adaptation des indices de sensibilité existants au contexte multidisciplinaire, dans un cas n'impliquant pas de couplages. Elle sera ensuite étendue aux couplages multidisciplinaires, où une revue bibliographique sur les approches de satisfaction des couplages sous incertitude orientera l'ajustement des stratégies d'estimation de sensibilité. Ensuite, l'étude explorera la création de nouveaux indices spécifiquement conçus pour l'optimisation multidisciplinaire. Des méthodes alternatives, comme le développement de métamodèles adaptés prenant en compte ces interactions complexes, seront également considérées.

Les méthodes élaborées seront appliquées à des cas analytiques de complexité variable pour valider leur efficacité et applicabilité, puis à des cas pratiques, notamment dans la conception de lanceurs tri-étages à l'ONERA (*c.f.* Figure 1) et de moteurs électriques à l'IFPEN.

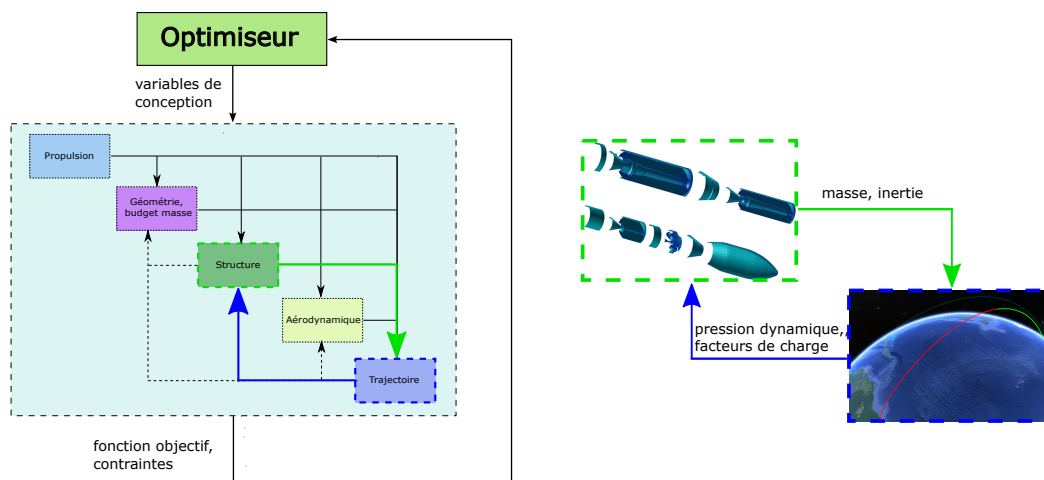


FIGURE 1 – A gauche : optimisation des variables de conception d'un lanceur via un système à 5 disciplines. A droite : détail des variables de couplage échangées entre les disciplines Structure et Trajectoire.

Properly constrained reference prior for a robust design of experiments in support of seismic fragility curves estimation

A. Van Biesbroeck^{†,1,2}, C. Gauchy^{§,3}, C. Feau^{§,2}, J. Garnier^{§,1}

[†] PhD student (presenting author). [§] PhD supervisors

PhD expected duration: Oct. 2022 – Sep. 2025

¹ CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, France
antoine.van-biesbroeck@polytechnique.edu

² Université Paris-Saclay, CEA, Service d'Études Mécaniques et Thermiques, France

³ Université Paris-Saclay, CEA, Service de Génie Logiciel pour la Simulation, France

Abstract

Seismic fragility curves of mechanical structures are key quantities of interest for probabilistic seismic risk assessment studies. They express the probability of failure of the mechanical structure conditional to a scalar value derived from the seismic ground motion, coined intensity measure.

This work targets studies for which the information on the response of the structure is limited to binary information of failure or not, for a given ground motion signal. We are particularly interested in experimental results obtained on complex systems, such as electrical devices, that cannot be modeled numerically, and for which a limited number of data are available.

For those, the Bayesian approach permits an efficient learning of the parameters which determine the fragility curves. It avoids the generation of unrealistic fragility curves such as unit step functions, which are common with classical methods. Nevertheless, it can still be affected by (i) the choice of the prior, whose influence is non-negligible in our study and has to be auditable in the nuclear industry; and by (ii) the occurrence of what we call degenerative scenarios. The latter are related to the distribution of the experimental results and can lead to improper posterior distributions, motivating even more the need of a robust choice for the prior.

We therefore propose a methodology based on a robust prior, constructed with the support of the ‘constrained’-reference-prior theory. This original contribution allows to build a prior which is both proper and objective. It serves the design of a planning of experiments based on the measure of the posterior’s sensitivity to the observed data, and supported by generalized sensitivity indices built with f -divergences.

The consistency of the method is evaluated on different case studies, and our results prove its capacity to efficiently solve the problem of degenerative sample generation with a limited number of data, enhancing any estimation of the seismic fragility curves.