

PROPOSITION DE STAGE EN COURS D'ETUDES

Référence : **DTIS-2026-02**
(à rappeler dans toute correspondance)

Lieu : Toulouse

Département/Dir./Serv. : DTIS/RIME

Tél. : 05.62.25.28.11

Responsable(s) du stage : Julien Demange-
Chryst, Jérôme Morio

Email : julien.demange-chryst@onera.fr
jerome.morio@onera.fr

DESCRIPTION DU STAGE

Thématique(s) : Mathématiques appliquées et leurs interactions, calcul scientifique

Type de stage : ☒ Fin d'études bac+5 ☒ Master 2 ☐ Bac+2 à bac+4 ☐ Autres

Intitulé : Prédiction conforme pour l'estimation d'évènements rares par apprentissage actif

Sujet : De nombreux systèmes physiques sont schématiquement décrits par une relation du type $Y = \phi(X)$, où l'entrée multidimensionnelle X est supposée aléatoire et où la sortie Y est déterminée via la fonction déterministe ϕ . Un exemple proéminent d'application est l'analyse d'un code de calcul boîte noire : ϕ représente alors un code de calcul, tel que des calculs de contraintes sur des structures mécaniques complexes et X les conditions extérieures dans lesquelles ce calcul est effectué. On peut notamment penser à un code de type éléments finis, dont la complexité rend impossible toute étude analytique de la fonction ϕ et donc de la sortie Y .

La défaillance de systèmes critiques peut engendrer des conséquences environnementales, économiques ou humaines potentiellement dramatiques. Il s'agit fort heureusement généralement d'évènements rares. Pour des raisons de sécurité et certifications, il est néanmoins crucial de quantifier le risque associé à de tels systèmes : il s'agit de l'analyse de fiabilité. Mathématiquement parlant, la défaillance d'un système est modélisée par un évènement de la forme $\{Y > t\}$, où t est un seuil de défaillance défini par des considérations physiques. Le risque associé au système est alors quantifié par la probabilité de défaillance $p_t = P(Y > t)$. Cependant, de par la complexité du code de calcul ϕ , p_t ne peut pas être calculer analytiquement.

L'estimation de la probabilité de défaillance est aujourd'hui un thème de recherche qui est largement étudié dans la littérature (méthodes Monte-Carlo, échantillonnage préférentiel [1], subset simulation [2]). Par ailleurs, le code de calcul ϕ étant potentiellement très coûteux en temps à évaluer, une famille répandue de méthodes d'estimation consiste à remplacer ϕ par un modèle de substitution, comme un processus gaussien par exemple (algorithme AK-MCS [3]). L'erreur d'approximation du méta-modèle est cependant difficile à maîtriser et ces méthodes ne permettent pas de quantifier correctement l'erreur d'estimation de p_t .

Dans ce contexte, l'objectif de ce stage est de construire un intervalle de confiance valide de la probabilité de défaillance, lorsqu'elle est estimée à l'aide d'un modèle de substitution, qui prend en compte l'erreur d'approximation de ϕ . Pour ce faire, une piste de recherche intéressante est la prédiction conforme [4]. Il s'agit d'une méthode statistique permettant de quantifier l'incertitude de n'importe quel modèle de prédiction en construisant des ensembles de prédiction statistiquement valides. Ce thème de recherche est aujourd'hui extrêmement dynamique dans la littérature.

Après une revue bibliographique sur l'analyse de fiabilité, les modèles de substitution et les méthodes de prédiction conforme, un premier axe de travail du stage serait, en se basant sur des travaux récents [5,6], d'appliquer une approche de prédiction conforme au méta-modèle utilisé et de la propager sur l'estimation de la probabilité de défaillance. Puis, un deuxième axe de travail serait de prendre en considération le caractère adaptatif des algorithmes de fiabilité, tel que l'algorithme AK-MCS. En effet, l'enrichissement actif du méta-modèle ne garantit plus l'hypothèse fondamentale d'échangeabilité des données. Des méthodes de prédiction conforme adaptées à cette situation seront à explorer [7] pour garantir la validité statistique de l'intervalle de confiance construit.

La méthode proposée durant le stage sera appliquée sur différents cas tests aérospatiaux.

- [1] J. A. Bucklew. Introduction to rare event simulation. *Springer Series in Statistics*. Springer-Verlag, New York, 2004.
- [2] F. Cérou, P. Del Moral, T. Furon, and A. Guyader. Sequential Monte Carlo for Rare Event Estimation, *Statistics and Computing*, 22(3):795-808, 2012.
- [3] Echard, B., Gayton, N., & Lemaire, M. AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation. *Structural safety*, 33(2), 145-154, 2011.
- [4] Angelopoulos, A. N., & Bates, S. ,A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *arXiv preprint arXiv:2107.07511*, 2021.
- [5] Jaber, E., Blot, V., Brunel, N. J., Chabridon, V., Remy, E., Iooss, B., & Leite, A. Conformal approach to gaussian process surrogate evaluation with marginal coverage guarantees. *Journal of Machine Learning for Modeling and Computing*, 6(3), 2025.
- [6] Pion, A., & Vazquez, E. Gaussian process interpolation with conformal prediction: methods and comparative analysis. In *International Conference on Machine Learning, Optimization, and Data Science*, 218-228, Springer Nature Switzerland, 2024.
- [7] Barber, R. F., Candes, E. J., Ramdas, A., & Tibshirani, R. J. Conformal prediction beyond exchangeability. *The Annals of Statistics*, 51(2), 816-845, 2023.

Est-il possible d'envisager un travail en binôme ? Non

Méthodes à mettre en œuvre :

- | | |
|---|--|
| ☒ Recherche théorique | ☐ Travail de synthèse |
| ☒ Recherche appliquée | ☐ Travail de documentation |
| ☐ Recherche expérimentale | ☐ Participation à une réalisation |

Possibilité de prolongation en thèse : Oui

Durée du stage : Minimum : 4 Maximum : 6

Période souhaitée : Début du stage au premier semestre 2026

PROFIL DU STAGIAIRE

Connaissances et niveau requis : Mathématiques appliquées, probabilités, statistique, maîtrise d'un langage de programmation.	Ecoles ou établissements souhaités : 3 ^{ème} année d'école d'ingénieur généraliste, Master 2.
--	---