www.onera.fr

ONERA
THE FRENCH AEROSPACE LAB

PROPOSITION DE SUJET DE THESE

Intitulé : Estimation du domaine probable d'atteinte d'une valeur cible modélisée par des simulateurs stochastiques à fidélités variables

Référence : **TIS-DTIS-2024-13**
(à rappeler dans toute correspondance)

Début de la thèse : Octobre 2024	**Date limite de candidature :** Mai 2024

Mots clés : Fiabilité, estimation de probabilités, quantification d'incertitudes, apprentissage actif

Profil et compétences recherchées

Mathématiques appliquées, probabilités / statistiques, méthodes numériques

Présentation du projet doctoral, contexte et objectif

Le contexte de cette thèse est la quantification d'incertitudes dans le cadre de simulations numériques coûteuses. En particulier, on s'intéresse à l'estimation d'un domaine de défaillance étant donné un certain nombre d'incertitudes. Ces incertitudes peuvent porter soit sur des variables en entrée du code de calcul, soit sur des variables internes au code. En effet, de nombreuses problématiques étudiées à l'ONERA se formalisent comme un problème de recherche d'un ensemble admissible de paramètres de contrôle pour lesquels une quantité d'intérêt respecte une certaine condition, par exemple de rester sous un seuil fixé. Il arrive que cette quantité d'intérêt ne soit pas déterministe pour un jeu de paramètres de contrôle mais stochastique par rapport à des variables d'environnement incertaines subies. Par exemple, dans le contexte aérospatial, la définition d'un domaine admissible de tir de missile robustes aux incertitudes (*e.g.*, fabrication, vol) à partir de modèle dynamique de référence est un enjeu essentiel. Dans le domaine de l'énergie, la conception d'éoliennes suppose la capacité à définir un domaine d'exclusion de dommage d'éoliennes qui est stochastique par rapport aux conditions de vent. Aussi, en général, afin de tenir compte des dispersions stochastiques, une estimation par méthodes de Monte-Carlo de la quantité d'intérêt est utilisée, impliquant ainsi un modèle de simulation stochastique dans la détermination d'un domaine admissible. Cependant, dans le contexte d'un modèle de simulation coûteux, des approches directes ne sont pas accessibles et donc une approche à base de *machine learning* (modèle de substitution aussi appelé métamodèle) doit être développée.

Considérons une quantité d'intérêt $f: \mathcal{D} \times \Xi \to \mathbb{R}$ qui dépend de variables de contrôle $\boldsymbol{x} \in \mathcal{D} \subset \mathbb{R}^d$ et de variables d'environnement incertaines $\boldsymbol{U} \in \Xi \subset \mathbb{R}^m$ modélisées par un vecteur aléatoire à densité de probabilité connue. Le problème peut donc s'écrire :

$$\Gamma = \{\boldsymbol{x} \in \mathcal{D} \mid g_L(\boldsymbol{x}) \leq s\}.$$

où $g_L(\boldsymbol{x}) = \mathcal{Q}_{1-\alpha}^{(L)}[f(\boldsymbol{x}, \boldsymbol{U})]$ est l'estimateur Monte Carlo du quantile d'ordre $1-\alpha$ basé sur un échantillon de taille n_L de la variable aléatoire $f(\boldsymbol{x}, \boldsymbol{U})$ et $s \in \mathbb{R}$ une valeur seuil définissant le domaine admissible.

On s'intéresse plus précisément au cas où les conditions aléatoires $\boldsymbol{U}$ sont intégrées dans la modélisation du système étudié. Les contraintes physiques sont alors obtenues par des simulateurs stochastiques boîtes noires, au sens où nous n'avons accès ni aux calculs internes au simulateur (trop complexe pour extraire des informations sur les variables de contrôle, code industriel propriétaire) ni au choix des réalisations du vecteur $\boldsymbol{U}$.

Dans le cadre de la thèse, l'objectif est de proposer une stratégie d'apprentissage adaptatif par *machine learning* d'un métamodèle du quantile g_L pour apprendre le domaine admissible. Cependant, dans le cas considéré g_L est le quantile d'une fonction coûteuse et son budget d'estimation est souvent prohibitif. Aussi, il est possible d'avoir recours à des simulateurs stochastiques avec différentes fidélités (différentes précisions et coûts de calculs), par exemple pour différentes tailles d'échantillons Monte Carlo. La stratégie d'apprentissage actif proposée devra alors permettre de sélectionner le niveau de précision sur le simulateur

stochastique g_L nécessaire pour un nouveau point en prenant en compte la proximité aux contours du domaine admissible.

Ainsi, la thèse devra proposer une stratégie d'apprentissage adaptatif d'un domaine admissible dans un cadre multi-fidélité tout en proposant des garanties sur l'estimation du domaine admissible en prenant compte les différentes sources d'incertitudes.

Les développements effectués dans cette thèse seront validés sur un cas d'application de la DGA, concernant le domaine de tir de missiles, et sur un cas de l'IFPEN dans le domaine de la conception mécanique robuste d'éoliennes offshore.

Le déroulement de la thèse suivra les étapes suivantes :

- Etat de l'art sur les techniques de détermination d'un domaine admissible avec deux focus : modèle de simulation stochastiques et multi-fidélité
- Développement d'une stratégie à base de Machine Learning et d'enrichissement adaptatif pour tenir compte de l'ensemble des sources d'incertitudes et de proposer des garanties sur l'estimation du domaine admissible
- Application sur des cas analytiques de complexité croissante
- Application sur les cas finaux DGA et IFPEN

Cette thèse est une collaboration entre l'INRIA, la DGA, l'IFPEN et l'ONERA.

Collaborations envisagées

INRIA, DGA, IFPEN et ONERA

Laboratoire d'accueil à l'ONERA	**Directeur de thèse**
Département : Traitement de l'information et systèmes	Nom : Régis Duvigneau
Lieu (centre ONERA) : Palaiseau	Laboratoire : INRIA Sophia Antipolis
Contact : Mathieu Balesdent et Loïc Brevault	Tél. :
Tél. : +33 1 80 38 66 08 / +33 1 80 38 66 88	Email : regis.duvigneau@inria.fr
Email : mathieu.balesdent@onera.fr / loic.brevault@onera.fr	

Pour plus d'informations : https://www.onera.fr/rejoindre-onera/la-formation-par-la-recherche