

Projet scientifique du Consortium

Problématique industrielle

Différentes industries se sont dotées, dans les dernières décennies, de codes de simulation numérique sophistiqués comme outils de prévision, d'optimisation et de décision (pour les moteurs thermiques et électriques, les systèmes de production dans le domaine de l'énergie nucléaire ou des énergies renouvelables comme l'éolien ou la géothermie, pour le stockage de gaz, pour la prévision des inondations etc.). Ces codes simulant des phénomènes complexes (comprenant des interactions, des non linéarités, des effets de seuil, etc.) sont de plus en plus réalistes, mais de plus en plus lourds en temps de calcul (parfois plusieurs jours).

L'exploitation de ces codes devient une problématique propre, en particulier lorsque l'objectif poursuivi nécessite un nombre important d'évaluations du code. Ceci est le cas dans les études de propagation d'incertitudes (estimation de l'incertitude des sorties du modèle) et les analyses de sensibilité (connaissance des entrées du modèle les plus influentes sur la variabilité des sorties) mais également dans les études d'optimisation, d'inversion ou de calibration en présence d'incertitudes.

Dans ce contexte, trois axes scientifiques listés ci-dessous apparaissent d'intérêt primordial et sont détaillés plus bas :

- Axe 1 : Calibration, validation et transposition de codes
- Axe 2 : Métamodélisation de codes dans des environnements complexes.
- Axe 3 : Optimisation et inversion en présence d'incertitudes, en particulier pour des distributions de probabilités en sortie de code.

Dans chacun de ces axes, une attention particulière sera portée à la grande dimension. Les problèmes réels font parfois intervenir plusieurs dizaines ou centaines d'entrées. Des avancées méthodologiques seront proposées pour tenir compte de cette difficulté supplémentaire.

Les travaux attendus du consortium se distinguent des recherches dominantes en machine learning par des spécificités liées à l'exploration de simulations numériques coûteuses. Cependant, Il apparait important de construire des passerelles entre les très nombreux développements récents en machine learning et le domaine de la simulation numérique.

Axe 1 : Calibration, validation et transposition de codes

1.1 Calibration

L'objectif est de proposer une méthodologie de calibration de code en abordant notamment la question de la planification d'expériences physiques réelles initiale et adaptative. Il s'agit aussi de développer des méthodes d'exploitation des résultats de ces expériences, par exemple en lien avec les techniques d'assimilation de données.

1.2 Validation de code

L'objectif est de modéliser l'erreur de modèle pour pouvoir la prendre en compte mais aussi de bien la comprendre, l'analyser pour pouvoir la réduire. Il s'agit également de définir de bons indicateurs pour comparer les données expérimentales et les simulations, notamment dans le cas de données complexes.

1.3 Transposition de codes

Dans certains cas, notamment dans le domaine nucléaire ou militaire, les expériences physiques ne peuvent être conduites à l'échelle réelle. Seule des expériences à petite échelle sont réalisables. L'objectif est de développer une méthodologie permettant la calibration d'un code dans le contexte de la transposition de petite à grande échelle.

Axe 2 : Métamodélisation pour des environnements "complexes"

2.1 - Variable mixtes et structure des entrées par arbre

Il s'agit de proposer des méthodologies pour la prise en compte de variables d'entrée mixtes ou de structure arborescente. Lorsque la valeur d'une variable d'entrée (ex : nombre de couches) crée de nouvelles variables (ex : l'épaisseur et le matériau de la nouvelle couche). Cela conduit à l'étude de métamodèles sur des variétés (e.g. les arbres).

2.2 Very small Data – données complexes, très riches (ex : sortie spatio-temporelle, film)

Certains codes de calcul prennent en entrée des variables temporelles ou spatiotemporelles, là encore les méthodologies doivent être adaptées. Enfin dans certains cas, les simulations coûtent très chères mais produisent des sorties très « riches », par exemple des films. Il s'agira donc de proposer des approches permettant d'exploiter la richesse des données produites. Une question connexe concerne la construction de métamodèles et la réduction de dimension sans stockage, « à la volée », les données étant trop volumineuses pour être stockées.

2.3 Réduction de modèle, boîtes grises, gradients

Il s'agit ici de proposer un métamodèle par réduction de modèle (base réduites, mélange de modèles structurés et non structurés avec quantification de la perte, réduction linéaire ou non linéaire). Un point important est de garder l'interprétabilité des modèles. L'idée de la métamodélisation de la boîte grise est d'aller exploiter les informations disponibles sur les équations du code (la physique) pour proposer un proxy plus rapide. Une autre approche pour proposer un métamodèle pertinent consiste à utiliser les dérivées qui peuvent être obtenues par différenciation automatique éventuellement.

2.4 Extrapolation

Cette question concerne la prédiction des métamodèles en extrapolation, c'est-à-dire en dehors du domaine de variation sur lequel il a été ajusté.

2.5 Multifidélités, agrégation de codes de calcul

Dans d'autres contextes, plusieurs codes de calculs compétitifs ou non peuvent contribuer à apporter de l'information. Il s'agit par exemple de simulateurs de complexité et précision différentes et à coûts de calcul associés différents. Des approches méthodologiques permettant d'intégrer ces divers codes de calcul doivent être proposées. Une question associée concerne la planification adaptative d'expériences adaptée aux simulateurs multiples multifidélités afin de tirer profit des simulations les moins coûteuses et utiliser les simulations les plus précises mais souvent les plus coûteuses à bon escient.

Axe 3 : Optimisation et inversion en présence d'incertitudes, en particulier pour des distributions de probabilités en sortie de code

Il s'agit d'explorer des codes de calculs qui soit prennent en entrées des variables incertaines ou qui soit sont des codes intrinsèquement stochastiques. Il s'agit ici de proposer des techniques de planification adaptative rapide pour l'optimisation et l'inversion avec prise en compte des incertitudes (ex : en grande probabilité). L'objectif de cet axe est également d'étudier l'optimisation et l'inversion pour des distributions de probabilité en sortie de code de calcul.