

Programme scientifique du consortium

Problématique industrielle

Au cours des dernières décennies, diverses industries ont adopté des codes de simulation numérique sophistiqués comme outils d'aide à la décision (pour la conception, la prévision et l'analyse de risque par exemple). Ces codes, utilisés dans des domaines tels que les moteurs thermiques et électriques, les procédés de fabrication, les systèmes de production d'énergie (nucléaire, éolien, géothermie), et la prévision des inondations, permettent de simuler des phénomènes complexes, intégrant interactions, non-linéarités et effets de seuil. En devenant de plus en plus réalistes, leur temps de calcul reste conséquent ou s'accroît, atteignant parfois plusieurs jours.

Conservier ces codes est essentiel car ils encapsulent une expertise considérable et représentent pour les entreprises un moyen de capitaliser leurs connaissances spécifiques. Cependant, pour en maximiser l'efficacité et en extraire le maximum d'informations, il est crucial de pouvoir leur associer des métamodèles, ou modèles de machine learning, c'est-à-dire des approximations rapides à évaluer. Ces métamodèles, construits de manière séquentielle ou non, peuvent ainsi alléger les charges de calcul et permettre des études plus poussées, notamment pour des études de propagation des incertitudes (estimation des incertitudes des sorties du modèle sous divers scénarios), des analyses de sensibilité (identification des entrées les plus influentes sur la variabilité des sorties). Ils sont également essentiels pour les études d'optimisation, d'inversion ou de calibration en présence d'incertitudes. L'association de codes de calcul et de métamodèles est une pratique de machine learning frugal.

De nos jours, la collecte de données est un processus quasi-systématique et ces dernières sont traitées par des modèles de plus en plus précis. Ainsi, l'intérêt croissant pour le couplage des modèles physiques et des modèles de données répond à un double objectif : accroître la performance des simulations tout en améliorant leur interprétabilité et leur explicabilité. En combinant la rigueur des modèles physiques avec la flexibilité des modèles de données, il est possible de créer des outils hybrides qui offrent à la fois des prédictions précises et une compréhension approfondie des phénomènes simulés. Ce couplage favorise une approche plus intégrée et plus globale, essentielle pour relever les défis actuels et futurs de la simulation numérique.

Les travaux précédents de CIROQUO ont apporté des avancées sur la gestion de nombreux aspects de l'analyse de simulateurs, permettant d'envisager des cas test présentant des combinaisons de challenges précédemment abordés et de nouveaux défis.

La gestion de la grande dimension, englobant l'analyse de sensibilité et la réduction de dimension, ainsi que la recherche d'explicabilité et d'interprétabilité des résultats occuperont un rôle central et transverse dans chacun des trois axes scientifiques majeurs sur lesquels se focaliseront les futurs travaux menés dans le cadre de CIROQUO, à savoir :

- Axe 1 : des approches "boîtes noires" aux approches "boîtes grises",
- Axe 2 : métamodélisation complexe,
- Axe 3 : enrichissement séquentiel pour l'optimisation et l'inversion.

L'ouverture aux avancées des domaines connexes concerne non seulement le machine learning, mais aussi les systèmes dynamiques, l'assimilation de données, le contrôle, ainsi que la recherche opérationnelle.

1 Des approches "boîtes noires" aux approches "boîtes grises"

Le premier axe de recherche s'intéresse à l'analyse des modèles "boîtes grises", c'est-à-dire aux modèles intégrant des informations partielles sur la modélisation du système étudié pour une meilleure interprétation et compréhension des phénomènes mis en jeu (par opposition aux modèles "boîtes noires")

complètement agnostiques), tout en conservant une certaine simplicité et efficacité par rapport aux modèles physiques complets (les équations mathématiques et les relations causales ne sont pas forcément toutes explicites).

1.1 Exploitation de la structure des codes

En premier lieu, nous aborderons les problématiques liées à la gestion de la structure interne de codes de simulation, en essayant de répondre aux trois problématiques suivantes.

1. (**Gestion des codes structurés**) Comment exploiter la structure interne d'un code de simulation (qu'il soit chaîné ou organisé sous forme d'arbre ou de graphe, faisant intervenir ou non plusieurs types de physiques) afin d'en accélérer l'étude, que ce soit pour de l'analyse de sensibilité, de la calibration, ou de la propagation d'incertitudes ?
2. (**Gestion des codes multi-échelles**) Comment intégrer de manière cohérente des simulations à différentes échelles temporelles et spatiales pour capturer des phénomènes complexes et réduire les incertitudes ?
3. (**Gestion des variables hiérarchiques (aussi appelées "méta-variables" ou "variables gachettes")**) Comment gérer, dans une perspective de meilleure interprétation des résultats et d'amélioration de la compréhension des phénomènes sous-jacents, les variables qui, lorsqu'elles atteignent certaines valeurs, entraînent des changements structurels dans le modèle de simulation ?

1.2 Exploitation des équations supportant les codes

Les modèles "**Physics-Informed**" représentent des avancées significatives dans le domaine de la modélisation prédictive en intégrant des connaissances physiques directement dans les modèles statistiques. Intégrer la physique dans les modèles de prédiction présente un intérêt majeur car cela permet de garantir que les solutions respectent les principes physiques sous-jacents, tout en utilisant efficacement les données observées pour ajuster les paramètres du modèle. Cette approche est cruciale dans les domaines où les données sont limitées ou bruitées, car elle renforce la robustesse et la généralisabilité des prédictions.

La construction de tels modèles pose néanmoins de nombreux défis, au niveau notamment de la formulation appropriée des contraintes physiques dans le cadre du modèle statistique, ou au niveau de la gestion des incertitudes et des biais introduits par ces contraintes.

Parmi ces modèles, les **PI-GPs** (Physics-informed Gaussian processes) sont particulièrement intéressants, car ils combinent la flexibilité des processus gaussiens avec la capacité d'incorporer des contraintes physiques, ce qui permet de modéliser des systèmes complexes tout en quantifiant les incertitudes de manière explicite. Les **PINNs** (Physics-informed Neural Networks) offrent quant-à-eux une généralisation plus flexible en terme de formulation, bien que plus difficilement interprétable et moins explicite dans la gestion des incertitudes. En utilisant des réseaux de neurones pour approximer les solutions tout en intégrant des équations physiques sous forme de contraintes, les PINNs sont utiles pour des problèmes où la structure physique exacte est complexe ou inconnue.

Dans le cadre de ce consortium, on propose ainsi d'étudier les atouts, les inconvénients mais également les complémentarités entre ces deux formulations, pour des applications où la flexibilité computationnelle est essentielle.

1.3 Analyse de la compensation d'erreur

La gestion de la compensation d'erreur dans les modèles de prédiction est un sujet central pour assurer la précision et la fiabilité des résultats. Dans le cadre du consortium CIROQUO, on propose d'aborder ce problème complexe à travers les quatre problématiques suivantes :

1. Comment **détecter** et évaluer efficacement les situations où des erreurs se compensent mutuellement dans différentes parties du modèle pour améliorer la qualité des prédictions ?
2. Quelles sont les conditions et les scénarios dans lesquels la compensation d'erreur peut être **bénéfique** ou **acceptable**, en conduisant à des prédictions plus robustes ou à une meilleure adaptation aux données ? Comment identifier les situations où la compensation d'erreur est **préjudiciable**, entraînant des prédictions biaisées ou une perte de précision significative ?

3. Comment développer des méthodes pour définir les zones de fonctionnement où les compensations d'erreur en mode prédiction sont compatibles avec celles observées sur les données d'apprentissage, en utilisant des techniques avancées de validation croisée et d'analyse comparative pour garantir la cohérence des résultats ?
4. Quelles stratégies peuvent être mises en oeuvre pour **réduire** les compensations d'erreur afin d'améliorer la fidélité des modèles de prédiction, incluant le recalibrage des modèles, l'amélioration de la qualité des données d'entrée, ainsi que l'intégration de techniques avancées de validation et de vérification ?

2 Métamodélisation complexe

Dans le cadre du consortium CIROQUO, l'axe de recherche sur la mise en place de méthodes de métamodélisation avancées vise à développer et à optimiser des techniques de modélisation permettant de simplifier et d'accélérer les simulations numériques tout en maintenant un haut niveau de précision, et s'organise autour de quatre sous-axes principaux.

2.1 Transposition, transfer-learning et multi-fidélité

Le premier thème porte sur le transfert efficace de modèles et de connaissances d'un domaine à un autre ou entre différentes échelles de fidélité afin d'améliorer la précision et de réduire les coûts de calcul. L'objectif est ainsi de développer des méthodes de transfert permettant de réutiliser des modèles préalablement entraînés sur des problèmes similaires ou des données de différentes résolutions, de mettre en place des techniques de modélisation multi-fidélité combinant des simulations de haute et basse fidélités pour optimiser les ressources computationnelles, et d'évaluer les performances de ces méthodes dans des contextes variés pour identifier les conditions sous lesquelles elles sont les plus efficaces.

2.2 Gestion des effets de seuil/falaise/bifurcation

Le second thème s'intéresse à modéliser et prédire de manière fiable les comportements non linéaires et les changements abrupts dans les réponses des systèmes complexes, tels que les effets de seuil ou falaises, et les bifurcations. Parmi les objectifs de cette partie sont inclus le développement de techniques de détection et de modélisation des points de bifurcation et des transitions de phase dans les simulations, l'intégration de méthodes d'analyse de sensibilité pour identifier les paramètres critiques influençant ces effets, et la mise en place d'approches robustes pour la quantification des incertitudes associées à ces phénomènes non linéaires.

2.3 Prédiction conforme / garanti (en probabilité) des prédictions

Le troisième thème se concentre sur la construction de prédictions fiables et si possible sur les garanties pour les modèles de simulation. L'objectif est cette fois de développer des méthodes de prédiction qui fournissent des intervalles de confiance probabilistes pour les prédictions (prédiction conforme), de garantir que ces prédictions respectent les contraintes physiques et statistiques des systèmes modélisés, et d'intégrer ces méthodes dans des outils de simulation pour une meilleure prise de décision.

2.4 Entrées et sorties complexes

Le dernier thème aborde le défi de la gestion des entrées et des sorties de simulation qui ne sont pas simplement des valeurs scalaires mais des vecteurs de grande taille ou des fonctions, voire des objets plus complexes comme des lois de probabilité ou des cartographies. Il s'agira ainsi de développer des techniques avancées pour la représentation et la manipulation de ces données, de mettre en place des méthodes permettant d'intégrer ces types de données dans les processus de modélisation et de simulation, et d'assurer que les modèles résultants soient capables de gérer efficacement ces entrées et sorties tout en maintenant une précision élevée et une interprétabilité claire.

3 Optimisation et inversion

Une fois un métamodèle défini, il est possible de l'améliorer pour des buts précis, que ce soit une amélioration globale de la qualité de prédiction ou plus généralement pour l'identification d'ensembles d'intérêt. Dans tous ces cas, une procédure séquentielle de sélection de nouvelles évaluations peut être adoptée, en parallèle ou non, via un ou plusieurs critères dédiés. Dans le cadre du consortium CIROQUO, quatre difficultés majeures pour la mise en place de telles procédures séquentielles seront étudiées.

On remarquera que l'ensemble des travaux listés dans cette section repose fortement sur l'utilisation de méthodes avancées d'optimisation, avec ou sans dérivées, que ce soit pour l'optimisation des hyperparamètres des métamodèles ou pour l'optimisation des critères séquentiels. Aussi, tout progrès dans ce domaine, susceptible de permettre l'obtention de modèles plus précis et l'utilisation de critères plus complexes, sera encouragé. On cherchera également, dans la mesure du possible, à rendre les développements, qui seront réalisés dans cet axe de recherche, compatibles avec des configurations présentant des entrées non scalaires, tels que des nuages de points, ou des variables catégorielles.

3.1 Optimisation bruitée / sous incertitudes

L'incertitude, qu'elle soit aléatoire (sur les observations ou les variables) ou épistémique (sur le modèle) doit être prise en compte dans la procédure d'optimisation et sa formalisation. Il s'agit donc de proposer un compromis entre budget disponible (e.g., calcul d'une espérance versus calcul d'un quantile extrême), précision (réduction d'incertitude adaptative), et adaptabilité (par exemple, pour des lois de probabilités en entrée et/ou sorties ou des scénarios).

3.2 Gestion des contraintes connues et cachées (ex : crashes)

Des contraintes sont généralement présentes dans les problèmes d'optimisation, liées à la paramétrisation du problème, à la méthode de résolution, aux sorties de la boîte noire, etc. Elles se présentent avec des complexités variées, allant de contraintes analytiques et évaluables instantanément à des contraintes cachées et difficilement prévisibles telles que des erreurs de maillage provoquant une erreur d'évaluation. Un axe de travail est donc d'exploiter au mieux une connaissance fine des contraintes, en l'intégrant dans la résolution, tout en palliant aux contraintes cachées.

3.3 Estimation de domaines admissibles et inversion

L'identification de parties du domaine où les valeurs des fonctions d'intérêt respectent certaines conditions est le pendant de l'optimisation. Une difficulté supplémentaire est la représentation de ces ensembles d'intérêt, en particulier lorsqu'il y a beaucoup d'entrées ou de sorties. Il s'agit en particulier de proposer des représentations à la fois compactes et représentative de ces ensembles pour les estimer efficacement.

3.4 Réduction de dimension pour l'optimisation

Sans hypothèse de structure sur le problème, l'optimisation en grande dimension se heurte au fléau de la dimension avec une augmentation exponentielle des budgets d'évaluations requis. Des méthodes de réduction de dimension, qui sélectionnent des variables ou des combinaisons de variables, permettent de traiter des problèmes avec un nombre important de variables. Néanmoins, l'articulation entre la réduction de dimension et l'optimisation reste à améliorer. D'une part, la réduction de dimension cherche à capter les variations alors que l'optimisation cherche des points de gradient nuls. D'autre part, représenter grossièrement les zones non pertinentes est suffisant alors qu'une certaine précision est requise pour identifier précisément un optimum.