



Sujet de thèse

Analyse statistique des résultats de simulations numériques de situations accidentelles sur les réacteurs à eau pressurisée

Contexte industriel

Le département Performance, risque industriel, surveillance pour la maintenance et l'exploitation (PRISME) participe à une mission essentielle de la R&D du groupe EDF : apporter un appui scientifique aux directions de production et d'ingénierie pour optimiser les performances des installations de production d'électricité, tout en garantissant leur sûreté.

Dans cette optique, PRISME porte le projet MOVIE dédié à la validation de modèles et au traitement des incertitudes en simulation numérique. En parallèle, CEA, EDF et Framatome (ex-AREVA) collaborent sur ce sujet au sein de l'Institut Tripartite (projet I3P « incertitudes ») pour réaliser des développements communs et asseoir des démarches et pratiques similaires dans le domaine de l'exploitation et des analyses de sûreté des réacteurs à eau pressurisée (REP).

Dans le cadre des démonstrations de sûreté nucléaire, l'ingénierie d'EDF utilise des outils de calcul scientifique dont la modélisation des phénomènes physiques à l'œuvre lors des transitoires accidentels est de plus en plus réaliste, mais de plus en plus difficile à appréhender et souvent plus lourde sur le plan informatique (temps CPU notamment). Dans le cas de l'étude des transitoires consécutifs à une perte accidentelle de réfrigérant primaire (APRP) par exemple, cela se traduit par l'emploi du code CATHARE2¹ en lieu et place des modèles simplifiés antérieurs et par le recours croissant à la mécanique des fluides numérique (CFD, *Computational Fluid Dynamics*, aussi appelée thermo-hydraulique locale). De telles évolutions sont motivées notamment par le vieillissement de certains composants des REP et par l'évolution des exigences réglementaires en matière de sûreté. Outre le problème plus aigu du budget de calcul (nombre limité de simulations numériques), les méthodes de prise en compte des incertitudes actuellement proposées aux ingénieurs (méthodes « BEPU » pour *Best Estimate Plus Uncertainty*) s'avèrent inadaptées aux indicateurs de risque d'intérêt (cas des analyses de sensibilité vis-à-vis de quantiles ou de super-quantiles) ou à des schémas de calcul dans lesquels interviennent des variables catégorielles ou dépendant du temps (ou de l'espace physique), parfois de manière peu régulière. De plus, elles ne prennent pas en compte la présence éventuelle de variables intermédiaires (cas d'un chaînage de codes). *Si des éléments de réponse sont disponibles dans la littérature sur ces sujets (voir ci-après), un travail de tri, d'adaptation, de mise en cohérence et de mise à l'épreuve de ces éléments est nécessaire, a minima.*

Cette thèse a pour objectif de mener ce travail méthodologique en lien direct avec le contexte d'étude industriel évoqué ci-avant. Trois objets d'étude, qui concernent la simulation numérique de scénarios d'APRP, sont proposés pour guider le travail de thèse :

¹ <http://www.irsn.fr/FR/Larecherche/outils-scientifiques/Codes-de-calcul/Pages/Le-code-CATHARE2-4661.aspx>

- A) Les transitoires de température des gaines du combustible nucléaire lors d'un APRP BI (brèche intermédiaire), simulés avec le code système CATHARE2. CATHARE2 est souvent qualifié de « *best estimate* » au sens où il fournit la meilleure modélisation envisageable actuellement pour la réalisation d'études industrielles. Le nombre de variables incertaines en entrée du modèle numérique est relativement grand : une trentaine sur un cas simplifié [17][7], plus de cent sur un cas réacteur. En revanche, le temps relativement court d'exécution du modèle permet d'envisager la réalisation de milliers de simulations.
- B) Les transitoires thermo-hydrauliques (pression, température et coefficient d'échange), simulés avec CATHARE2, qui sont considérés dans l'évaluation du risque de rupture brutale de la cuve par choc froid et les facteurs de marge correspondants (avec prise en compte du phénomène de pré-chargement à chaud), évalués par le code de mécanique Cuve1D [17], chaîné à CATHARE2. Comme dans le cas A), les transitoires thermo-hydrauliques peuvent présenter des sauts (fortes variations quasi-instantanées).
- C) Le comportement neutronique du cœur (risque de retour en criticité) dans un scénario de dilution hétérogène inhérente dont la simulation implique la transmission par Code_Saturne (CFD) du champ 3D de la concentration en bore au code de neutronique COCAGNE. Ces codes sont qualifiés de codes « *best effort* » au sens où ils modélisent aussi finement que possible les phénomènes physiques. Le nombre de calculs CFD est fortement limité (quelques dizaines voire moins). Par ailleurs, certaines variables d'entrée du calcul neutronique, qui concernent l'état du réacteur nucléaire, sont discrètes (position des grappes).

L'objectif général de la thèse est la définition de méthodes d'analyse des résultats de simulations numériques de ces transitoires accidentels pour l'identification des principales sources d'incertitude vis-à-vis du risque considéré et de configurations pénalisantes. Dans la pratique, il s'agirait notamment de conforter l'identification des sources et des configurations qui est réalisée par la méthode PIRT (*Phenomenon Identification Ranking Table*) [8]. Les méthodes envisagées proviennent du domaine de l'apprentissage statistique [9].

État de l'art et verrous scientifiques

Dans cet objectif, le doctorant travaillera à lever les quatre verrous qui suivent.

- 1) **Analyse statistique d'échantillons de transitoires ou de champ physique (de grande dimension).** Certaines grandeurs physiques d'entrée ou de sortie des simulations numériques ont une nature fonctionnelle, par exemple un champ spatial 3D ou une quantité dépendant du temps. En pratique, les données correspondantes sont des vecteurs de grande dimension. Un axe de travail porte sur l'exploitation de méthodes de réduction de dimension [12] pour l'analyse d'échantillons de telles données, préalablement à la mise en œuvre d'autres méthodes d'apprentissage statistique visant i) à détecter des données « aberrantes » (*outliers*), ii) à opérer une classification des données ou iii) à construire un métamodèle². L'intérêt de i) et de ii) est d'asseoir ou d'améliorer la compréhension des phénomènes physiques les plus importants pour les grandeurs d'intérêt par une expertise physique des différentes classes ou *outliers* distingués. Pour ce faire, les travaux antérieurs [1] pourraient être prolongés par l'emploi de nouvelles approches faisant appel à la notion statistique de profondeur [13][25] ou de déformation (*warping*) [4], issue du domaine de la reconnaissance de formes. Dans le

² Un métamodèle est un modèle de substitution capable d'approcher suffisamment précisément le résultat d'une évaluation du modèle numérique considéré pour un coût et un temps de calcul faibles. Il est construit et validé à l'aide d'un certain nombre de calculs réalisés avec le modèle numérique initial (à l'aide d'un plan d'expérience). Construire un métamodèle n'est généralement pas une fin en soi mais une solution à la mise en œuvre d'algorithmes (estimations statistiques, optimisation, etc.) inapplicable directement au modèle initial.

cas des transitoires thermo-hydrauliques simulés par CATHARE2 dont les sauts posent problème [1], une approche plus fastidieuse consisterait à aider les procédures d'apprentissage en complétant les données des transitoires avec des informations sur l'origine des sauts constatés, ce qui supposerait une analyse humaine de chaque transitoire simulé (à défaut d'une modification de CATHARE2). Dans le cas d'un chaînage de codes, le lien entre analyse des données fonctionnelles intermédiaires et les grandeurs d'intérêt évaluées en bout de chaîne [19] est également un sujet à approfondir.

- 2) **Analyse de sensibilité (AS) ciblée et conditionnelle.** Les méthodes usuelles de criblage (méthode de Morris) ou d'analyse de sensibilité (décomposition de la variance) identifient les facteurs et phénomènes les plus influents sur la variabilité globale de la grandeur d'intérêt considérée. Cette pratique peut être améliorée par la considération des simulations les plus pertinentes pour l'objectif de sûreté. L'idée est d'utiliser et d'approfondir les techniques d'AS ciblée et conditionnelle développées récemment [21].

Plusieurs références relatives aux deux autres sujets présentés ci-après sont basées sur une méthode de construction de métamodèle par modélisation par processus gaussien [7], encore appelée krigeage. Il s'agit d'un cadre de travail flexible et puissant qui sera a priori privilégié dans les travaux sur ces sujets.

- 3) **Développement de méthodes de traitement d'incertitudes dédiées aux chaînes de calcul.** La simulation des accidents nucléaires fait souvent intervenir des modèles numériques chaînés avec des variables intermédiaires de nature fonctionnelle (cas d'application A) et B)). La question de la prise en compte explicite de ces dernières dans les méthodes d'apprentissage statistique se pose pour exploiter de manière optimale un budget de calcul en traitant l'information apportée par les variables intermédiaires et en ouvrant la voie à la définition de stratégies de planification et d'allocation de calculs qui tirent parti du fait qu'un modèle numérique (la thermo-hydraulique en l'occurrence) soit beaucoup plus coûteux sur le plan computationnel qu'un autre modèle de la chaîne (lien à étudier avec [14]). Le travail proposé porte sur l'analyse de sensibilité [3] et la construction de métamodèle [16].
- 4) **Détermination automatique de scénarios accidentels pénalisants.** L'une des premières étapes dans une analyse de sûreté est de déterminer un ensemble de scénarios dits « pénalisants ». Cette étape, souvent réalisée manuellement sur la base d'un PIRT, gagnerait à être automatisée via l'utilisation de techniques d'apprentissage statistique. L'élaboration de techniques d'optimisation pour la détermination d'un pire cas [12][23], ou d'inversion d'un domaine de défaillance [5], peut nécessiter l'utilisation de métamodèles, comme le krigeage. Dans le contexte des simulations accidentelles, le krigeage apparaît délicat à mettre en œuvre du fait de discontinuités inhérentes aux phénomènes modélisés [2] ou de la présence de variables non continues [20][24].

Programme de travail

première année	deuxième année	troisième année
<p><i>Principaux sujets à traiter : 1) et 2)</i></p> <ul style="list-style-type: none"> — Revue bibliographique des sujets 1 et 2 (3 mois) — Appropriation des cas d'application A) et B) — Définition et évaluation de plans d'expériences (PEX) CATHARE2 (cas A) et B)) — Étude de la possibilité de tirer parti d'informations complémentaires sur l'origine des sauts — Analyse des résultats des PEX <ul style="list-style-type: none"> - Analyse de sensibilité/criblage avec un grand nombre de variables - Classification des transitoires et lien avec les entrées incertaines 	<p><i>Principaux sujets à traiter : 2) et 3)</i></p> <ul style="list-style-type: none"> — Approfondissement du sujet 2) : analyse de sensibilité ciblée en petite dimension — Revue bibliographique des sujets 3) et 4) (3 mois) — Mise en place du cas d'application C) (informatique, formalisation du problème) — Évaluation d'un premier plan d'expériences sur le cas C) — Définition d'une méthodologie globale en réponse aux sujets 3) et 4) et évaluation de cette méthodologie sur des modèles symboliques jouets 	<p><i>Principaux sujets à traiter : 3) et 4)</i></p> <ul style="list-style-type: none"> — Amélioration de la méthodologie en réponse aux sujets 3) et 4) — Mise en œuvre de la méthodologie proposée sur le cas B) — Mise en œuvre de la méthodologie proposée sur le cas C) — Rédaction du mémoire (4 mois) — Préparation de la soutenance (1 mois)

Encadrement du doctorant

Cette thèse sera réalisée au sein du département PRISME de EDF R&D à Chatou et sera co-encadrée par Mathieu Couplet (docteur) et Bertrand Iooss (habilité à diriger des recherches) qui portent la compétence en traitement des incertitudes. Roman Sueur, qui travaille également à PRISME, apportera des appuis techniques pour le cas d'application B) et scientifiques pour le verrou 1) essentiellement. Un apport du projet « Marges Déterministes » (départements MFEE et PERICLES d'EDF R&D) est attendu pour la prise en main du cas d'application C).

Étant données les compétences des autres personnes qui interviennent dans l'encadrement de la thèse, mais surtout le caractère fortement applicatif du sujet de thèse, l'encadrement académique sera porté par Nathalie Marie (CEA Cadarache), experte dans les modèles physiques utilisés dans les simulations d'accidents thermo-hydrauliques, au sein de l'école doctorale I-MEP2 (<http://www.grenoble-inp.fr/ecole-doctorale-i-mep2-448443.kjsp>) de l'Institut Polytechnique de Grenoble.

Un co-encadrement sera porté dans le cadre du projet I3P « incertitudes » par Amandine Marrel du CEA Cadarache (habilité à diriger des recherches).

Références

- [1] B. Auder (2011). *Classification et modélisation de sorties fonctionnelles de codes de calcul*. Thèse de l'université Paris VI.
- [2] J. Baccou (2017). *Contributions à l'analyse de données spatialisées et applications dans les études de risque*. Mémoire d'habilitation à diriger des recherches de l'université Aix-Marseille.
- [3] Y. Caniou (2012). *Global sensitivity analysis for nested and multiscale modelling*. Thèse de l'université Blaise Pascal de Clermont-Ferrand.
- [4] G. Charpiat, O. Faugeras, R. Keriven, P. Maurel (2015). *Approximations of Shape Metrics and Applications to Shape Warping and Empirical Shape Statistics*. Statistics and Analysis of Shapes, Krim & Yezzi eds, Birkhäuser, pp. 363-395.
- [5] C. Chevalier (2013). *Fast uncertainty reduction strategies relying on Gaussian process models*. Thèse de l'université de Bern (Suisse).
- [6] T. Delage, B. Iooss, M. Couplet, A. Marrel, H. Geiser, A.-L. Popelin (2017). *Propositions d'outils statistiques pour les analyses d'incertitudes en thermohydraulique système - Application à la méthode CATHSBI*. Rapport interne EDF R&D, référence 6125-3120-2017-04519-FR.
- [7] T. Delage, R. Sueur and B. Iooss (2018). *Robustness analysis of epistemic uncertainties propagation studies in LOCA assessment thermal-hydraulic model*. ANS Best Estimate Plus Uncertainty International Conference (BEPU 2018), soumis.
- [8] M.J. Griffiths, J.P. Schlegel, T. Hibiki, M. Ishii, I. Kinoshita, Y. Yoshida (2014). Phenomena identification and ranking table for thermal-hydraulic phenomena during a small-break LOCA with loss of high pressure injection. *Progress in Nuclear Energy* 73, pp. 51-63.
- [9] T. Hastie, R. Tibshirani, J. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction (second edition)*. Springer Series in Statistics.
- [10] B. Iooss, A. Marrel (2017). *An efficient methodology for the analysis and modeling of computer experiments with large number of inputs*. Proceedings of UNCECOMP 2017, 2nd ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering.
- [11] B. Iooss and A. Marrel (2018). *Advanced methodology for uncertainty propagation in computer experiments with large number of inputs*. ANS Best Estimate Plus Uncertainty International Conference (BEPU 2018), soumis.
- [12] D. Jones, M. Schonlau, W. Welch (1998). *Efficient Global Optimization of Expensive Black-Box Functions*. *Journal of Global Optimization* 13, 455–492.
- [13] J. Lee, M. Verleysen (2007). *Nonlinear Dimensionality Reduction*. Springer-Verlag New York.
- [14] L. Le Gratiet, C. Cannamela (2015). *Cokriging-Based Sequential Design Strategies Using Fast Cross-Validation Techniques for Multi-Fidelity Computer Codes*. *Technometrics*, 57(3), pp. 418-427.
- [15] S. López-Pintado, J. Romo (2007). *Depth-based inference for functional data*. *Computational Statistics & Data Analysis* 51, pp. 4957-4968.
- [16] S. Marque-Pucheu, G. Perrin, J. Garnier (2017). *Efficient sequential experimental design for surrogate modeling of nested codes*. Disponible à <https://hal.archives-ouvertes.fr/hal-01657827>.
- [17] A. Martin et al. (2008). *Engineering approach for the Reactor Pressure Vessel Integrity in PTS conditions*. Proceedings of the ASME Pressure Vessels and Piping conference.

- [18] P. Mazgaj, J.-L. Vacher, S. Carnevali (2016). *Comparison of CATHARE results with the experimental results of cold leg intermediate break LOCA obtained during ROSA-2/LSTF test 7*. EPJ Nuclear Sci. Technol. 2(1).
- [19] S. Nanty (2015). *Quantification des incertitudes et analyse de sensibilité pour codes de calcul à entrées fonctionnelles et dépendantes*. Thèse de l'université Grenoble Alpes.
- [20] P.Z.G. Qian, H. Wu, C.F.J. Wu (2008). *Gaussian Process Models for Computer Experiments With Qualitative and Quantitative Factors*. Technometrics, 50(3).
- [21] H. Raguét, A. Marrel (2018). Target and Conditional Sensitivity Analysis with Emphasis on Dependence Measures. Rapport du CEA DEN, disponible à arxiv.org/abs/1801.10047.
- [22] C.E. Rasmussen, C.K.I. Williams (2006). *Gaussian Processes for Machine Learning*. The MIT Press, disponible à <http://www.gaussianprocess.org/gpml/>.
- [23] N. Srinivas, A. Krause, S. Kakade, M. Seeger (2010). *Gaussian process optimization in the bandit setting: no regret and experimental design*. ICML'10 Proceedings of the 27th International Conference on International Conference on Machine Learning, pp. 1015-1022.
- [24] Q. Zhou, P.Z.G. Qian, S. Zhou (2011). *A Simple Approach to Emulation for Computer Models With Qualitative and Quantitative Factors*. Technometrics, 53(3), pp. 266-273.
- [25] Y. Zuo et R. Serfling (2000). *General notions of statistical depth function*. The Annals of Statistics 28(2), pp. 461–482.