

BILAN SCIENTIFIQUE du Projet **DICE**

Le consortium **DICE** « *Deep Inside Computer Experiments* » est un projet de recherche qui a réuni universitaires, industriels et institutions publiques pour avancer sur les thèmes de l'exploration et de l'optimisation des grands codes de calcul (mot clé : *computer experiments*). Il a été piloté par le département 3MI, laboratoire commun à l'École des Mines de Saint-Étienne (ENSM-SE) et ARMINES. Ce projet de recherche a regroupé des partenaires industriels ou institutionnels (Total, Renault, EDF, IRSN et ONERA) et des participants académiques (ENSM-SE, UJF Grenoble, Univ. Aix-Marseille, Univ. Orsay). Le projet a débuté fin 2006 et a duré 3 ans.

Site web : <http://dice.emse.fr/>

Les codes de calcul simulant des phénomènes variés profitent en permanence du développement des capacités de calcul des processeurs. Mais ce gain en réalisme se paye en difficulté à appréhender le comportement du code développé ceci pour deux raisons concourantes. Tout d'abord, le temps calcul d'un code peut rapidement être prohibitif, en dépassant la journée, ce qui rend son exploration délicate. D'autre part, les performances des codes se rapprochant de la réalité, leur compréhension est tout aussi délicate que celle du réel, avec de très nombreux paramètres en jeu dont les effets sont difficiles à appréhender.

Ainsi, de nombreuses questions simples d'énoncé restent sans réponse satisfaisante et partagée, comme la détection de paramètres influents ou l'influence de la méconnaissance de certains paramètres sur les prévisions effectuées. Le consortium DICE s'est penché sur l'étude approfondie de deux types d'objets, les *métamodèles* et les *plans d'expériences numériques*, déclinés principalement au travers des thèmes de recherche suivants : métamodèles et optimisation, krigeage bayésien et propagation d'incertitudes, planification d'expériences numériques, krigeage et autres métamodèles, criblage et analyse de sensibilité.

RESULTATS MAJEURS, FAITS MARQUANTS

De nombreuses contributions méthodologiques ont résulté du consortium DICE, et sont détaillées ci-dessous. Parmi les plus marquantes, on peut citer une méthode stochastique de planification d'expériences permettant la génération d'expériences dans des domaines contraints, la parallélisation de la méthode EGO pour l'optimisation de grands codes de calculs, et la mise au point d'une méthode adaptative pour le calcul de probabilité de défaillance. Les méthodes produites ont été testées et validées sur des cas d'étude multiples proposés par les partenaires industriels.

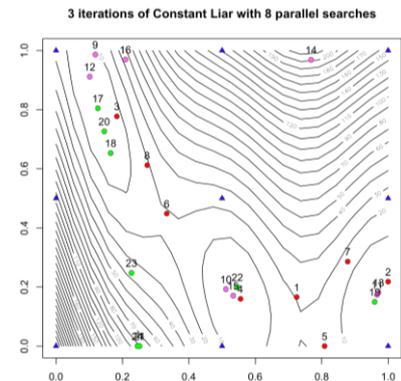
En terme de production scientifique, on retiendra : 4 livrables logiciels, plus de 30 livrables/présentations à usage interne (280 Mo !), 3 thèses soutenues, 2 thèses démarrées pendant le consortium, 8 publications, et de nombreuses communications dans des conférences. Enfin, fort des compétences acquises sur les *computer experiments*, un colloque européen sur ce thème a été organisé à l'ENSM-SE en partenariat avec le réseau de statisticiens ENBIS. Il a été l'occasion de riches interactions entre les participants de DICE et les autres équipes de recherche européennes travaillant sur le sujet. Site web : <http://www.emse.fr/enbis-emse2009/>.

BILAN DÉTAILLÉ

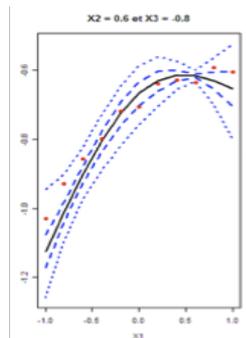
Métamodèles et optimisation

L'utilisation de métamodèles pour l'optimisation de grands codes de calcul est rendue nécessaire par le temps de calcul prohibitif d'une simulation. Cependant, l'efficacité d'une optimisation basée sur un métamodèle passe par la prise en compte de l'erreur de modélisation, ce qui explique la préférence pour un métamodèle de type krigeage. Les travaux réalisés ont consisté à adapter l'algorithme EGO (Efficient Global Optimization) pour le calcul parallèle, ce qui présente l'avantage de pouvoir accélérer considérablement la durée totale de l'optimisation.

Production scientifique. Package R «*DiceOptim*», thèse de D. Ginsbourger (parties I et III), 1 livrable, 2 publications ([3], [4]).



Krigeage bayésien et propagation d'incertitudes



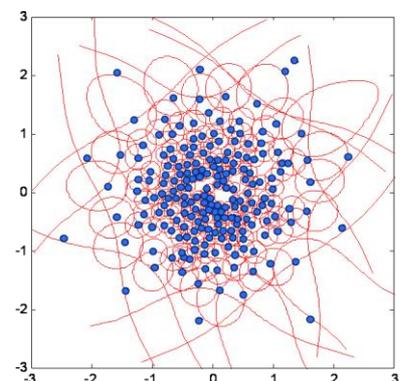
Lorsqu'un métamodèle est utilisé pour la propagation d'incertitudes, l'erreur d'estimation s'ajoute aux sources d'incertitude. Concernant le krigeage, le krigeage universel (KU) et les méthodes bayésiennes existantes n'intègrent souvent que l'erreur d'estimation d'un sous-ensemble des paramètres. Les travaux réalisés dans DICE ont consisté à fournir une interprétation rigoureuse des formules du KU, et à développer la méthodologie permettant de réaliser une étude bayésienne complète.

Production scientifique. 1 livrable, 1 publication ([5]).

Planification d'expériences numériques

1. Construction de plans SFD en phase exploratoire

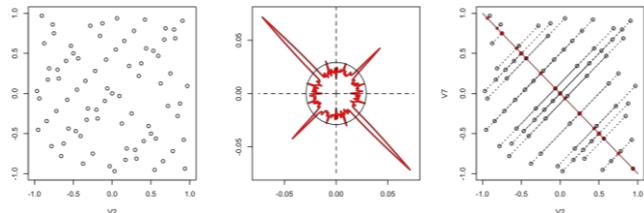
Un grand code de calcul modélise un phénomène souvent très complexe. Pour cette raison, lorsqu'on cherche à définir un premier plan d'expériences numériques (phase exploratoire), il est souhaitable de remplir le plus possible le domaine : ce sont les *Space-Filling Designs* (SFD). Les travaux réalisés dans DICE ont conduit à proposer une nouvelle méthode de construction de plans SFD, de nature stochastique, basée sur le processus ponctuel de Strauss. La souplesse de cette approche permet de tenir compte des caractéristiques particulières souvent rencontrés dans la pratique : domaines contraints, distributions des entrées.



Planification d'expériences numériques (suite)

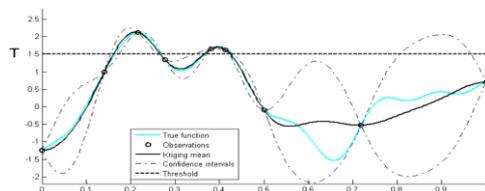
2. Evaluation des plans SFD de la phase exploratoire

Comme premier plan d'expériences, les plans SFDs sont bien adaptés pour la modélisation de phénomènes complexes, mais se comportent souvent assez mal en projection sur des sous-espaces de plus petite dimension. Or cette contrainte apparaît naturellement dans une démarche de réduction de la dimensionnalité des variables d'entrée. Les travaux réalisés dans DICE ont abouti à définir un critère qui permet de quantifier la qualité de plans vis-à-vis de leur comportement en projection. Un tri peut donc être effectué a priori parmi les nombreux plans SFDs candidats.



3. Plans adaptatifs pour l'estimation d'une probabilité de défaillance

Après la phase d'exploration du code de calcul, il est souhaitable de planifier les expériences futures selon une démarche adaptative, personnalisée à l'objectif recherché. Dans le cadre de la fiabilité, par exemple, le métamodèle se doit d'être fidèle non pas globalement mais localement sur la frontière définissant la contrainte de défaillance. Les travaux réalisés dans DICE ont abouti à définir une démarche adaptative qui permet d'atteindre cet objectif. Basée sur un métamodèle de krigeage, elle utilise la mesure de l'erreur de modèle pour réaliser le compromis nécessaire entre recherche locale et globale.

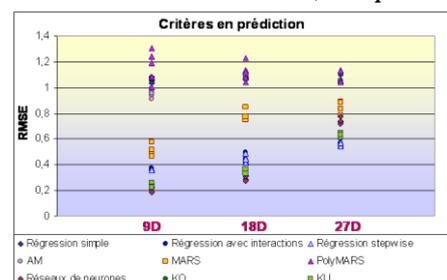


Production scientifique. Package R « DiceDesign », thèse de J. Franco, thèse de V. Picheny (chapitre 5), 1 livrable, 3 publications ([1], [6], [8]).

Krigeage et autres métamodèles

1. Evaluation de métamodèles

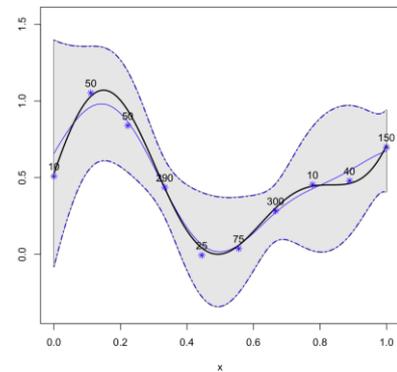
Les métamodèles usuels sont constitués de modèles déterministes définis par des familles de fonctions comme les polynômes, polynômes trigonométriques, fonctions cassures type polyMARS, réseaux de neurones, etc. ou des modèles probabilistes comme le krigeage. Ces derniers sont bien adaptés pour fournir un ordre de grandeur de l'erreur de modèle, ce qui est d'un intérêt particulier ici. Les travaux réalisés dans DICE ont consisté d'une part à un travail d'implémentation (krigeage, outils pour la validation) et d'autre part à l'étude de problèmes industriels variés. En particulier, la comparaison du krigeage avec les autres types de métamodèles a montré l'adaptabilité de la méthode à diverses situations, en en fait donc un métamodèle privilégié.



Krigeage et autres métamodèles (suite)

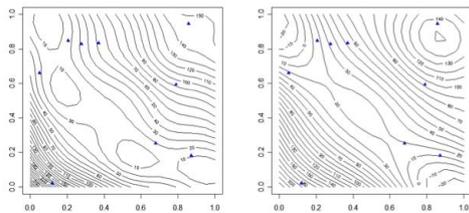
2. Krigeage pour les simulateurs stochastiques

Le krigeage, habituellement utilisé pour les simulateurs déterministes, peut être adapté pour les simulateurs stochastiques. Dans ce contexte, de nouvelles questions se posent. Par exemple, s'agissant de simulations de Monte Carlo, la précision d'un calcul est réglable par l'utilisateur. Se pose alors la question de la répartition de l'effort de simulation : à budget fixé, est-il préférable de faire un grand nombre de simulations imprécises ou bien quelques simulations très précises ? Le travail réalisé dans DICE a permis d'apporter des réponses à cette question, après un travail préliminaire de formalisation et d'implémentation.



3. Krigeage et contraintes

Le dernier volet - et non le moindre ! - au sujet des métamodèles de krigeage consiste à modifier les noyaux de covariance pour tenir compte des spécificités des problèmes étudiés. La première application traitée dans DICE a ainsi permis de construire un modèle de krigeage intégrant les symétries naturellement présentes dans un problème. D'autres applications, basées sur la théorie des espaces de Hilbert à noyau reproduisant (RKHS), permettent de généraliser le krigeage à des contraintes continues.

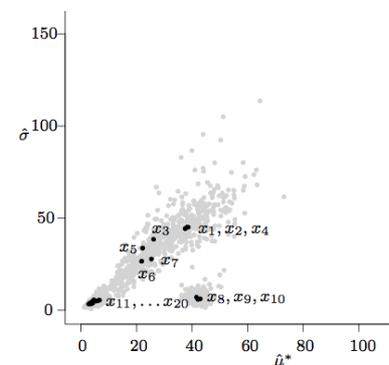


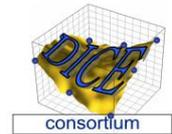
Production scientifique. Packages R « *DiceEval* », « *DiceKriging* », thèse D. Ginsbourger (partie II), thèse V. Picheny (chapitre 6), thèse B. Gauthier (en cours), thèse N. Durrande (en cours), 2 livrables, 1 publication ([2]).

Criblage et analyse de sensibilité

Le criblage ou screening consiste à sélectionner les facteurs réellement influents sur une variable d'intérêt. L'analyse de sensibilité permet de quantifier l'influence des facteurs et leurs interactions. Dans DICE, un état de l'art des différentes méthodes existantes a été effectué, aboutissant à une étude comparative sur des cas industriels. Une méthode originale d'analyse de sensibilité a également été proposée.

Production scientifique. 2 livrables, 1 publication ([7]).





PRODUCTION SCIENTIFIQUE

Logiciels - Packages R

4 packages en langage R ont été publiés sur CRAN, le site officiel de R, et sont maintenant téléchargeables :

- [DiceDesign](#)
- [DiceEval](#)
- [DiceKriging](#)
- [DiceOptim](#)

Livrables à usage interne

Plus de 30 livrables/présentations ont été produits.

Thèses de Doctorat

- Franco J. (2008), *Planification d'expériences numériques en phase exploratoire pour la simulation des phénomènes complexes*, thèse CIFRE Total. [[télécharger](#)]
- Ginsbourger D. (2009), *Multiplés métamodèles pour l'approximation et l'optimisation de fonctions numériques multivariées*, thèse de l'ENSM-SE. [[télécharger](#)]
- Picheny V. (2009), *Improving accuracy and compensating for uncertainty in surrogate modeling*, thèse de l'ENSM-SE en cotutelle avec l'University of Florida. [[télécharger](#)]
- Gauthier B. (2007-). Thèse de l'ENSM-SE en cours.
- Durrande N. (2008-). Thèse de l'ENSM-SE en cours.

Publications

1. Franco J., Bay X., Dupuy D. et Corre B. (2008), Planification d'expériences numériques à partir du processus ponctuel de Strauss. Soumis. [[preprint sur HAL](#)].
2. Ginsbourger D., Dupuy D., Badea A., Roustant O. et Carraro L. (2009), A note on the choice and the estimation of Kriging models for the analysis of deterministic computer experiments, *Applied Stochastic Models for Business and Industry*, **25** (2), 115-131. [[résumé](#)]
3. Ginsbourger D., Helbert C. et Carraro L. (2008), Discrete mixtures of kernels for Kriging-based optimization, *Quality and Reliability Engineering International*, **24** (6), 681-691. [[résumé](#)]
4. Ginsbourger, D., Le Riche, R. and Carraro L. (2010), Kriging is well-suited to parallelize optimization. In *Computational Intelligence in Expensive Optimization Problems - Adaptation Learning and Optimization*, pages 131-162, Springer Berlin Heidelberg. [[résumé](#)] [[preprint sur HAL](#)]
5. Helbert C., Dupuy D. et Carraro L. (2009), Assessment of uncertainty in computer experiments : from Universal to Bayesian Kriging, *Applied Stochastic Models for Business and Industry*, **25** (2), 99-113. [[résumé](#)]
6. Picheny V., Ginsbourger D., Roustant O., Haftka R.T. et Kim, N.-H., Adaptive Designs of Experiments for Accurate Approximation of Target Regions, *Journal of Mechanical Design*, **132** (7), 071008 (9 pages). [[résumé](#)]
7. Pujol G. (2009), Simplex-based screening designs for estimating metamodels, *Reliability Engineering and System Safety*, **94**, 1156-1160. [[résumé](#)][[preprint sur HAL](#)].
8. Roustant O., Franco J., Carraro L. et Jourdan A. (2010), A radial scanning statistic for selecting space-filling designs in computer experiments, in A. Giovagnoli, A.C. Atkinson, B. Thorsney and C. May, "MODA - 9 - Advances in Model-Oriented Design and Analysis", Springer (Physica-Verlag), p. 189-196. [[transparents](#)][[software](#)].

Communications orales

Les travaux de DICE ont fait l'objet de nombreuses communications orales, non répertoriées dans ce document.