



---

Université Blaise Pascal – Clermont II  
Institut Français de Mécanique Avancée

---

Cécile MATTRAND

E-mail : [cmattrand@ifma.fr](mailto:cmattrand@ifma.fr)

Directeur de thèse : Pr. Maurice Lemaire, Co-encadrant scientifique : Dr. Jean-Marc Bourinet

Laboratoire de Mécanique et Ingénieries (EA 3867 - FR TIMS / CNRS 2856)  
Université Blaise Pascal et Institut Français de Mécanique Avancée  
Campus de Clermont-Ferrand - Les Cézeaux BP 265  
63175 Aubière Cedex

## **CALCUL A COUT REDUIT DES INDICES DE SOBOL' PAR APPRENTISSAGE SVM**

Cette recherche apporte une contribution au domaine de l'analyse de sensibilité globale en développant une stratégie économique de calcul des indices de Sobol'. La solution s'appuie sur une méthode d'apprentissage statistique basée sur les *Support Vector Machines* (SVM). La méthodologie développée est appliquée et validée sur quelques exemples relevant de la mécanique.

### **Analyse de sensibilité globale :**

Les méthodes d'analyse de sensibilité globale sont des méthodes quantitatives qui permettent de mesurer la contribution de chaque variable aléatoire (ou groupe de variables aléatoires) d'entrée sur la sortie d'un modèle numérique traduisant un comportement physique, ce qui est particulièrement utile pour identifier les paramètres potentiellement influents. Parmi l'ensemble des méthodes proposées (Saltelli *et al.*, 2000), certaines comme les indices de Sobol' (Sobol', 1993) évaluent le poids de chaque variable (ou groupes de variables) sur la réponse du modèle via une analyse basée sur la variance. Les indices de Sobol' étudient la répercussion de la variabilité des variables aléatoires d'entrée, supposées indépendantes, sur la variabilité de la réponse en mesurant la réduction de la variance de la sortie d'un modèle lorsqu'une ou plusieurs variables aléatoires sont fixées. Formulés à l'aide de variances et d'espérances conditionnelles, ils sont classiquement estimés à l'aide des simulations de Monte Carlo ou de méthodes de génération de points à discrétance faible, telles que les suites de Sobol' randomisées.

### **Problématique :**

L'application à des exemples analytiques simples a montré que le calcul des indices de Sobol' nécessite un nombre important d'appels au modèle ce qui rend leur obtention difficile pour des modèles numériques supposés gros consommateurs de ressources informatiques. Le travail présenté vise à définir une stratégie économique de calcul des sensibilités.

### **Solution proposée :**

L'originalité de la solution proposée réside dans la construction d'un substitut au modèle, cette approximation étant obtenue par une régression basée sur les *Support Vector Machines* (Smola & Schölkopf, 2003), d'où la notation : SVR (Support Vector Regression). Ce substitut est construit dans l'espace des variables Gaussiennes centrées réduites. Les indices de Sobol' sont ensuite estimés via le substitut SVR, qui ne nécessite alors plus d'appels au modèle.

Le principe de régression par *Support Vector Machines* repose sur l'algorithme suivant : après une phase d'apprentissage sur  $n$  couples (données-réponse), le substitut SVR calcule la réponse associée à une nouvelle donnée. Le modèle initial est alors approché par une fonction simple, ici basée sur les SVM. La difficulté dans la construction du substitut SVR réside dans le choix et le nombre  $n$  de couples (données, réponse), formant ce que l'on appelle la base d'apprentissage, ainsi que dans le choix des hyper-paramètres du modèle.

Dans la solution proposée, une procédure de sélection systématique des hyper-paramètres est mise en place. Elle repose notamment sur les connaissances du problème a priori et sur le principe de validation croisée.

Deux méthodes déterministes de génération de la base d'apprentissage sont testées. La première méthode consiste à placer les points d'apprentissage uniformément dans un hyper-cube grâce aux suites de Sobol' déterministes (Bratley & Fox, 1988), puis à passer ensuite dans un espace de variables Gaussiennes centrées réduites où le substitut au modèle est construit. La deuxième méthode consiste à placer les points de la base d'apprentissage de manière la plus uniforme possible dans une hyper-sphère de rayon choisi dans l'espace des variables Gaussiennes centrées réduites.

### Exemples traités :

L'efficacité de la stratégie proposée en termes du nombre d'appels au modèle numérique est évaluée sur un exemple analytique (Nie & Ellingwood 2004). Le modèle est défini par :

$$f(\mathbf{x}) = f(x_1, x_2, x_3) = -\left(x_1^2 + x_2^2 + x_3^2 - 2x_1x_2 - 2x_2x_3 - 2x_3x_1\right)/2 - (x_1 + x_2 + x_3)/\sqrt{3} + 3 \quad (1)$$

où  $x_1$ ,  $x_2$  et  $x_3$  sont des variables aléatoires Gaussiennes centrées réduites et indépendantes.

La construction d'un substitut SVR au modèle à partir d'une base d'apprentissage de 500 couples données-réponse permet d'obtenir des indices de Sobol' (premiers ordres, ordres supérieurs et indices totaux) de même précision relative que ceux obtenus classiquement avec des suites de Sobol' randomisées avec 4 500 000 appels au modèle, soit ici un gain total de l'ordre de 10000 en termes de nombre d'appels au modèle numérique.

La stratégie de calcul à coût réduit est également testée sur le problème industriel PROSIR ou *Round Robins on Probabilistic Approach for Structural Integrity of Reactor Pressure Vessel* (PROSIR, 2006). Le problème consiste en une étude thermo-mécanique de la résistance d'une cuve face à un choc thermique pressurisé suite à une injection d'eau de refroidissement. Le modèle utilisé permet de déterminer la nocivité d'un défaut sous le revêtement, en tenant compte de l'effet de l'irradiation sur les propriétés à rupture des matériaux. L'étude menée permet d'obtenir les sensibilités du problème mécanique aux variables d'entrée. Là encore, la méthodologie SVR proposée met en évidence le gain en coût de calcul, par rapport à une approche classique basée sur une utilisation des suites de Sobol' randomisées.

### Bibliographie :

Bratley, P. and Fox, B. L. 1988. ALGORITHM 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on mathematical software* 14(1): 88-100

Nie, J. and Ellingwood, B.R. 2004. A new directional simulation method for system reliability, part II: application of neural networks. *Probabilistic Engineering Mechanics* 19: 437-447.

PROSIR. 2006. Probabilistic structural integrity of PWR reactor pressure vessel. In *Round Robins on Probabilistic Approach for Structural Integrity of Reactor Pressure Vessel*. OECD-NEA PROSIR Workshop, Lyon, France

Saltelli, A, Chan, K and Scott, E. M. 2000. Sensitivity analysis. *Wiley*.

Smola, A. and Schölkopf, B. 2003. A tutorial on support vector regression. *Technical Report*.

Sobol', I. M. 1993. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments* 26: 1-42.