



Sujet de stage – année scolaire 2018/2019

Etude d'un estimateur d'indice de sensibilité par perturbation de lois

CONTEXTE :

Au sein d'EDF R&D, le département PRISME a pour mission de proposer des solutions innovantes pour une exploitation plus performante des différents moyens de production du groupe EDF : maîtrise des risques, amélioration de la sûreté, optimisation des performances technico-économique, maîtrise de la durée de vie. Il contribue via un projet de recherche interne aux développements et à la diffusion de méthodes de traitement des incertitudes dans les outils de calcul scientifique, et notamment des méthodes d'analyse de sensibilité (AS) qui visent à classer des variables incertaines en fonction de leur impact sur une quantité d'intérêt évaluée numériquement (ici un quantile $q^\alpha = \inf \{t \in \mathbb{R} \mid F_Y(t) \geq \alpha\}$). De manière encore plus générale, l'analyse de sensibilité peut être vue comme une méthode d'exploration de modèle, permettant de synthétiser dans un indice quantitatif (un scalaire S_i ou une courbe) l'influence globale d'une ou plusieurs variables aléatoires X_i sur la distribution de $Y = g(X)$ (où g est un code numérique).

OBJECTIF DU STAGE :

Le stage portera sur l'utilisation de nouveaux indices de sensibilité, développés dans le cadre d'un travail de thèse¹ au sein d'EDF. Appelés PLI (*Perturbed Law based sensitivity Indices*)². La technique consiste à « perturber » un paramètre de la loi d'une des entrées $F_{i\delta}$, afin d'observer dans quelle mesure cette perturbation modifie le quantile de sortie $q_{i\delta}^\alpha = \inf \{t \in \mathbb{R} \mid F_{Y_{i\delta}}(t) \geq \alpha\}$. Bien que récemment introduit, les indices PLI présente un grand intérêt industriel. L'objectif du stage est d'enrichir le corpus méthodologique disponible autour de ces outils. Il s'articulera selon plusieurs phases (qui pourront se chevaucher) :

- 1) Un travail bibliographique qui fera la synthèse des éléments disponibles au sujet des PLI.
- 2) Un travail théorique et applicatif sur les propriétés de l'estimateur des indices PLI (convergence en loi de processus³, estimation par bootstrap ou noyaux).
- 3) Une étape importante de la méthode consiste à transformer le problème dans un espace standardisé (l'ensemble des distributions des paramètres d'entrées y sont des Gaussiennes). Une partie du travail sera de caractériser le lien entre la perturbation dans l'espace standard et son influence dans l'espace physique.
- 4) L'implémentation en langage Python de l'estimation statistique de ces indices à partir d'un échantillon de calculs, mais également d'un test du caractère significatif de l'impact de chaque variable incertaine sur la quantité d'intérêt. De ce fait d'autres méthodes d'AS pourront par ailleurs être identifiées et étudiées⁴.

D'autres questions sont encore ouvertes et pourront être également étudié lors du stage suivant les affinités théoriques ou applicatives du candidat. On citera par exemple l'adaptation des indices PLI aux quantiles $T_{\alpha,\beta}$ ⁵ (quantile au seuil α , avec une confiance β), ou la perturbation de plusieurs paramètres d'entrées simultanément.

PROFIL :

Etudiants de M2 probabilités / statistiques ou grandes écoles.

COMPETENCES SOUHAITEES :

Bases solides en statistiques et en mathématiques appliquées.

Aisance en informatique.

ENVIRONNEMENT INFORMATIQUE :

Linux, Latex, Python, logiciel R.

CONTACTS :

jerome.stenger@edf.fr

bertrand.iooss@edf.fr

DUREE ENVISAGEE :

5 - 6 mois à partir de mars 2019

LIEU :

EDF R&D

Département Performance, Risque Industriel et Surveillance pour la Maintenance et l'Exploitation

6, Quai Watier - 78401 Chatou

¹ <http://www.theses.fr/2014BORD0061>

² <https://hal.inria.fr/hal-01552361/document>

³ <https://www.math.u-psud.fr/~meliot/master/poisson.pdf>

⁴ Indices de Sobol, indices HSIC.

⁵ http://openturns.github.io/openturns/master/examples/data_analysis/quantile_estimation_wilks.html



Internship proposal – year 2018/2019

Study of the estimator of a perturbed law based sensitivity indices

CONTEXT:

The purpose of the PRISME department at EDF R&D is to conceive innovative solutions for a more performant exploitation of the different means of production: risks management, surety improvement, optimization of performance, durability control. The department, through an internal research project, works on the development of methods for uncertainty quantification, in particular on methods of sensitivity analysis (SA) whose purpose is to assess how the uncertainty in the output of a system (here a quantile $q^\alpha = \inf \{t \in \mathbb{R} \mid F_Y(t) \geq \alpha\}$) can be apportioned to different sources of uncertainty in its inputs. More generally, sensitivity analysis can be seen as an exploratory method of a model, reduced to a simple quantity (typically a scalar S_i or a curve), in order to determine how changes in one random variables X_i can affect the distribution of an outcome $Y = g(X)$ (where g is a numerical code).

OBJECTIVE OF THE INTERNSHIP:

The internship will focus on a new sensitivity index, developed during a Phd¹ at EDF, called PLI (*Perturbed Law based sensitivity Indices*)². The principle is to assess the influence of a perturbation on a parameter of the input distribution $F_{i\delta}$, on the output quantity, the quantile of a safety margin $q_{i\delta}^\alpha = \inf \{t \in \mathbb{R} \mid F_{Yi\delta}(t) \geq \alpha\}$. While newly introduced, the PLI indices has shown a great deal of interest for Industrial application. The goal of the internship is to expand the tools and methodologies available with this technique. The internship will be structured in several stages (that might overlap):

- 1) A bibliographic study, in order to synthesize the actual state of the PLI knowledge.
- 2) A theoretical and applicative work on the properties of the PLI indices estimator (Convergence of Stochastic Processes³, bootstrapping or kernel estimation)
- 3) The technique relies on a transformation of the problem into a standard space (where every input distributions are Gaussian). One of the student work will be to evaluate the impact of a perturbation in the standard space onto the physical space.
- 4) An implementation of the statistical estimation of the PLI indices will be realized on Python. Other SA methods⁴ will be identified and studied in order to compare the quality of the quantity evaluated.

Many other questions remain open and might be investigated during the internship, regarding the affinity of the student towards theoretical or applicative work. One can for instance mention the adaptation of the PLI indices to the quantiles $T_{\alpha,\beta}$ ⁵ (quantile at threshold α , with the confidence β), or the perturbation of the density of several inputs simultaneously.

PROFILE:

Students of a Master in probability/statistic or engineering school.

COMPETENCES DESIRED:

Solid basis in statistics and applied mathematics.

Ease with computational work.

COMPUTING ENVIRONMENT:

Linux, Latex, Python, R.

CONTACTS :

jerome.stenger@edf.fr

bertrand.iooss@edf.fr

INTERNSHIP TIME:

5 - 6 months from March 2019

LOCATION:

EDF R&D

Département Performance, Risque Industriel et Surveillance pour la Maintenance et l'Exploitation

6, Quai Watier - 78401 Chatou

¹ <http://www.theses.fr/2014BORD0061>

² <https://hal.inria.fr/hal-01552361/document>

³ <https://www.math.u-psud.fr/~meliot/master/poisson.pdf>

⁴ Sobol indices, HSIC indices.

⁵ http://openturns.github.io/openturns/master/examples/data_analysis/quantile_estimation_wilks.html