



# Offre de Thèse : Nouvelles stratégies pour la conception robuste de procédés

Sensibilité globale, apprentissage automatique et logique floue

## English (New strategies for robust process design)

### Project Description (Global Sensibility, Machine Learning and Fuzzy logic)

During classical process design steps, taking inaccuracies and/or uncertainties into account on the data is carried out either by local sensitivity studies (one parameter at a time) or by calculation of flexibility indices. The methods are achieved via simulation and technical-economic evaluation models, or through meta-models (multi-parameter approach, emulator, “surrogate model”, etc.).

In recent years, new global sensitivity methods (Sobol method, Polynomial chaos, etc.) have emerged in numerical simulation, allowing uncertainty propagation modeling. Similarly, Machine Learning methods have been established with considerable expansion in dealing with problems of classification, prediction, etc. Finally, possibility theory better known as fuzzy logic is a good mathematical tool particularly well suited for handling flexibly, incomplete information, the unsharpness of classes of objects or situations, or the gradualness of sensitivity criterion. The aim of the thesis is part of a disciplinary triptych: Mathematical Optimization, Machine Learning, and Process Engineering. The

research project will investigate cutting-edge machine learning for application in process design. One of the key project challenges is to propose a technique for process control to handle dynamic situations whereby process conditions can change in a “random” way to specific changes in operation. The proposed disciplinary triptych may offer improved process design-control steps over classical control loops.

The thesis aims to renew the strategies used for process design through the use - separated or combined - of Global Sensitivity techniques, Machine Learning and Possibility Theory. It is a part of a generic approach to meet the following scientific challenges:

- The problem positioning relative to the state of the art,
- The formulation of robust design problems based on global sensitivity methods,
- The use of the three techniques (above) to reach more “robust” solutions. In particular cases, it may allow taking into account inaccuracies/uncertainties inherent in the design stage.
- The hybridization of methods and comparison of the proposed approach.

## Candidate’s profile and skills

- The targeted Ph.D. candidate shall have (or expect to achieve) an MSc and/or engineer degree with a strong background in control and computer engineering, signal processing and/or in applied mathematics and statistics, or a comparable field of study.
- Besides a strong theoretical background, computer skills, including the first experience in using state-of-the-art machine learning and/or Fuzzy logic frameworks and programming environment (e.g., Julia, Matlab, python, R, C++), will be particularly expected.
- Very high motivation for solving methodological problems and willingness to work on Engineering Process Design.
- An interest to work in an interdisciplinary team.

## Hosting research team

We offer the opportunity of a PhD thesis at the **Toulouse INP-ENCIASET** Engineering School at the **LGC** (Chemical Engineering Laboratory). The LGC is supported by its three tutorships, the CNRS, the National Polytechnic Institute of Toulouse and the Paul Sabatier University of Toulouse. Through its six scientific divisions, the 300 staff of LGC address five major societal challenges: water and effluents, energy, bio-refinery, materials processing and health engineering.

The successful candidate will form part of a multidisciplinary research team in the Process and System Engineering department (**PSI**).

## Application modalities

Candidates are requested to send a detailed CV, a motivation letter and at least one recommendation letter from a reference researcher to Prof. FLOQUET Pascal and OUARET Rachid :

`pascal.floquet@ensiacet.fr`

`rachid.ouaret@toulouse-inp.fr`

## En Français (et détails)

### Résumé

Lors des étapes de conception “classique” de procédés, via des modèles de simulation et d’évaluation technico-économique, ou au travers de métamodèle (approche multi-paramétrique, émulateur, “surrogate model”,...), la prise en compte des imprécisions et/ou incertitudes sur les données est effectuée par des études de sensibilité locale (un paramètre à la fois) ou le calcul d’indices de flexibilité.

Ces dernières années, de nouvelles méthodes de **sensibilité globale** (méthode de Sobol, polynômes de chaos,...) et leur implantation efficace ont vu le jour, permettant la modélisation et la propagation d’incertitudes en simulation numérique. De même, les méthodes d’apprentissage automatique (“machine learning”) ont connu une expansion considérable, dans le traitement des problèmes de partitionnement ou de discrimination, par exemple. Enfin, la modélisation de données mal connues reste un atout majeur de la théorie des possibilités, plus connue sous le nom de logique floue. La problématique de la thèse s’inscrit dans un triptyque discipline : le Génie des Procédés, l’Optimisation Mathématiques et l’Apprentissage Automatique.

La thèse vise à renouveler les stratégies utilisées pour la conception des procédés par l’utilisation - séparée ou combinée - de techniques de sensibilité globale, d’apprentissage automatique et la théorie des possibilités. Elle inscrit dans une démarche générique pour répondre aux défis scientifiques suivant (i) positionnement de la problématique par rapport à l’état de l’art, (ii) formulation des problèmes de conception robuste en s’adossant sur les méthodes de sensibilité globale, (iii) l’utilisation des trois techniques (ci-dessus) afin d’aboutir à des solutions plus “robustes” - permettant de prendre en compte des imprécisions/incertitudes inhérentes à l’étape de design - et plus “flexibles”, c’est-à-dire capables de supporter, sans une trop forte perte d’efficacité, des aléas lors de la vie du procédé, et (iv) hybridation des méthodes et comparaison des démarches proposées.

## Mots clés

Sensibilité globale, Polynômes de chaos et Méthode de Sobol, Propagation d'incertitude, Flexibilité et Robustesse, Apprentissage Automatique et Logique floue.

## Profile candidat, pré-requis et compétences recherchées

### Formation initiale :

De formation Bac +5 minimum, le(la) candidat(e) doit être diplômé(e) d'un Master ou d'un diplôme d'Ingénieur à forte composante en automatique/traitement de signal ou en mathématique appliquée et statistique.

### Compétences techniques recommandées et souhaitables :

- Une bonne maîtrise des méthodes "classiques" de l'apprentissage automatique est indispensable.
- Une connaissance informatique en termes de programmation est nécessaire.
- Maîtriser les outils d'aide à la décision.
- Une première expérience dans un des domaines suivant :
  - la segmentation floue (théorie des possibilités),
  - l'analyse de la propagation des incertitudes,
  - conception des procédés,est un grand atout
- Connaissances en gestion industrielle appliquée aux procédés seraient un plus.
- La **motivation** et l'**envie d'apprendre** seront des atouts non négligeables.

## Localisation

Le poste occupé sera localisé à INP-ENSIACET au sein du Laboratoire de Génie Chimique (**LGC**) et rattaché au département Procédé et Système Industriels (PSI). Le laboratoire est une unité de recherche placée sous la triple tutelle du CNRS, de l'Institut National Polytechnique de Toulouse et de l'Université Paul Sabatier.

Le(a) candidat(e) sera inscrit(e) à l'École Doctorale de Mécanique, Énergétique, Génie civil & Procédés (**MEGeP**) de l'Institut National Polytechnique de Toulouse (INPT) (Établissements co-accrédités : Université Paul Sabatier, INSA, ISAE, IMT Mines Albi).

## Candidature

Sur dossier et entretien. Votre candidature devra comporter :

- un CV
- une lettre de motivation.
- le rapport du stage de Master 2 (si disponible et si possible).
- Les résultats de classement en Master 2 ou de l'année  $n - 1$  si celles du Master 2 ne sont pas disponibles.
- Toute autre pièce jointe que vous jugerez utile.

Votre candidature doit être transmise par voie électronique à FLOQUET Pascal et OUARET Rachid par courriel :

**pascal.floquet@ensiacet.fr**

**rachid.ouaret@toulouse-inp.fr**

## Description détaillée

### Contexte

L'efficacité des systèmes industriels ainsi que leurs performances nécessite une conception optimisée des procédés. L'étape de la conception est fondamentale et est en soi une tâche complexe. La conception "classique" de procédés est souvent élaborée via des modèles de simulation et d'évaluation technico-économique, en posant la question de la prise en compte des imprécisions et des incertitudes des modèles. En même temps, un intérêt particulier a été porté au cours de ces dernières années sur le développement des méthodes de propagation d'incertitude pour une optimisation robuste des procédés [10].

Ces incertitudes sont prise en compte que ce soit sur (i) les données de base (production, qualité des produits, ...) qui correspondent à un choix opéré aux étapes de spécification du projet ou sur (ii) les données technico-économiques (durée d'investissement, taux de crédit, ...).

Ces choix obèrent les capacités de l'installation à fonctionner sous d'autres régimes, que ceux-ci soient souhaités (augmentation de production) ou subis (changement de cadres financiers, lors de la réalisation du projet). Qui plus est, l'utilisation, dans l'étape de conception du procédé, d'un métamodèle (par exemple en utilisant une approche "surrogate model"), qui est une démarche de plus en plus courante en génie des procédés, nécessite la prise en compte des imprécisions déjà mentionnées, mais également celles des paramètres de ces modèles simplifiés, souvent établis via des méthodes de type Monte-Carlo. Ceci est actuellement effectué par des études de sensibilité locale ou le calcul d'indices de flexibilité/robustesse/résilience [5, 11].

Les études de sensibilité locale ont pour objectif d'obtenir une conception d'unités plus "robuste" en faisant varier les paramètres d'intérêt, un à la fois et les uns à la suite des autres. Elles s'apparentent peu ou prou à la méthodologie HAZOP, qu'on peut rencontrer en étude de sûreté, ou à la réalisation de scénario (optimiste, pessimiste, neutre,...), qui se retrouve dans les études prospectives. Largement employées, elles souffrent de deux défauts majeurs : la non-prise en compte de l'interaction des paramètres, par le fait même d'étudier les paramètres les uns indépendamment des autres et l'aspect "local" de l'indicateur, autour du "design nominal" déjà calculé. De ce fait il s'agit plus d'une vérification de la flexibilité de la conception, effectuée a posteriori, qu'une méthode de conception "robuste" prenant en compte les incertitudes/imprécisions et aboutissant à la "meilleure" conception possible.

Le calcul d'indices, de flexibilité ou de résilience, a été largement utilisé depuis le milieu des années 80 et l'article pionnier de Swaney et Grossmann [19]. Une récente recension des travaux sur ce sujet [6] a permis de dénombrer au moins 5 indices différents : (i) indices de flexibilité de Swaney et Grossmann [19], (ii) de Pitsikopoulos et Mazzuchi [12], (iii) de Dimitradis et Pitsikopoulos [4], (iv) indice de résilience de Saboo et Morari [13], (v) indice de flexibilité "volumétrique" de Lai et Hui [8].

Sans rentrer dans une discussion détaillée, ces divers indicateurs diffèrent par la prise en compte des incertitudes/imprécisions (stochastique vs déterministe) et l'ensemble des activités englobées (conception "simple" des unités, conception des unités et du système de contrôle-commande,...). Ces démarches pèchent cependant toutes par la non différenciation des notions d'incertain et d'imprécis, de manque d'information et de variabilité. Et leur approches, basées essentiellement sur des méthodes d'optimisation, nécessitent une formulation complexe et des temps de calcul importants pour des résultats n'étant garantis que "localement" optimaux...

## Objectifs scientifiques

Ces dernières années, de nouvelles méthodes d'analyse de sensibilité, dite "**globale**" [18] (méthode de Sobol [16, 17, 7], polynômes de chaos [9, 3],...) et leur implantation efficace ont vu le jour, notamment sous l'impulsion de la Communauté Européenne (équipes des Pr Andrea Saltelli et Stefano Tarantola, Joint Research Centre, European Commission, Institute for the Protection and Security of the Citizen, Ispre (Italie), par exemple)[14], permettant la modélisation et la propagation d'incertitudes dans tous les domaines de la simulation numérique. De même, les méthodes d'apprentissage automatique ("machine learning") ont connu une expansion considérable, dans le traitement des problèmes de partitionnement ou de discrimination, par exemple [5, 15, 1]. Enfin, l'utilisation conjointe, en modélisation, de données mal connues (via la théorie des possibilités) et de données imprécises (connues par leur distribution de probabilité, par exemple) a été rendue possible, par les travaux de Baudrit, Couso et Dubois [2].

Les étapes de la démarche proposée sont les suivantes :

1. L'analyse des atouts et des faiblesses des méthodologies actuelles de conception robuste de procédés. Analyse des outils actuels d'analyse de sensibilité, "machine learning" et de théorie des possibilités.

2. La formulation des problèmes de conception robuste, selon le type des données imprécises et incertaines, au travers d'exemples de la littérature ou du LGC (conception de réseaux d'échangeurs de chaleur, bio-raffinerie,...).
3. L'hybridation des méthodes et comparaison des démarches proposées et plus anciennes, au crible de critères tels que la facilité de modélisation, la représentativité de l'approche, le temps calcul, ...

## References

- [1] Yun Bai, Zhenzhong Sun, Jun Deng, Lin Li, Jianyu Long, and Chuan Li. Manufacturing quality prediction using intelligent learning approaches: A comparative study. *Sustainability*, 10(1):85, 2018.
- [2] Cédric Baudrit, Inés Couso, and Didier Dubois. Joint propagation of probability and possibility in risk analysis: Towards a formal framework. *International Journal of Approximate Reasoning*, 45(1):82–105, 2007.
- [3] Jean-Marc Bourinet. Form sensitivities to distribution parameters with the nataf transformation. In *Risk and Reliability Analysis: Theory and Applications*, pages 277–302. Springer, 2017.
- [4] Veniamin D Dimitriadis and Efstratios N Pistikopoulos. Flexibility analysis of dynamic systems. *Industrial & Engineering Chemistry Research*, 34(12):4451–4462, 1995.
- [5] Pascal Floquet, Gilles Hétreux, Raphaelé Hétreux, and Lucille Payet. Analysis of operational heat exchanger network robustness via interval arithmetic. In *Computer Aided Chemical Engineering*, volume 38, pages 1401–1406. Elsevier, 2016.
- [6] Ignacio E Grossmann, Bruno A Calfa, and Pablo Garcia-Herreros. Evolution of concepts and models for quantifying resiliency and flexibility of chemical processes. *Computers & Chemical Engineering*, 70:22–34, 2014.
- [7] S Kucherenko et al. A new derivative based importance criterion for groups of variables and its link with the global sensitivity indices. *Computer Physics Communications*, 181(7):1212–1217, 2010.
- [8] Sau M Lai and Chi-Wai Hui. Process flexibility for multivariable systems. *Industrial & Engineering Chemistry Research*, 47(12):4170–4183, 2008.
- [9] Andre Nataf. Determination des distribution dont les marges sont donnees. *Comptes Rendus de l Academie des Sciences*, 225:42–43, 1962.
- [10] Omid Nejadseyfi, Hubert Geijselaers, and Ton van den Boogaard. Robust optimization based on analytical evaluation of uncertainty propagation. *Engineering Optimization*, 51(9):1581–1603, 2019.

- [11] Lucille Payet, Raphaële Thery Hétreux, Gilles Hétreux, Florent Bourgeois, and Pascal Floquet. Flexibility assessment of heat exchanger networks: from a thorough data extraction to robustness evaluation. *Chemical Engineering Research and Design*, 131:571–583, 2018.
- [12] EN Pistikopoulos and TA Mazzuchi. A novel flexibility analysis approach for processes with stochastic parameters. *Computers & Chemical Engineering*, 14(9):991–1000, 1990.
- [13] Alok K Saboo, Manfred Morari, and Duncan C Woodcock. Design of resilient processing plants –viii. a resilience index for heat exchanger networks. *Chemical Engineering Science*, 40(8):1553–1565, 1985.
- [14] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008.
- [15] Ahmed Shokry, Sergio Medina-González, and Antonio Espuña. Mixed-integer multiparametric approach based on machine learning techniques. In *Computer Aided Chemical Engineering*, volume 40, pages 451–456. Elsevier, 2017.
- [16] Ilya M Sobol. Sensitivity estimates for nonlinear mathematical models. *Mathematical modelling and computational experiments*, 1(4):407–414, 1993.
- [17] Ilya M Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and computers in simulation*, 55(1-3):271–280, 2001.
- [18] Bruno Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety*, 93(7):964–979, 2008.
- [19] Ross Edward Swaney and Ignacio E Grossmann. An index for operational flexibility in chemical process design. part i: Formulation and theory. *AIChE Journal*, 31(4):621–630, 1985.