

OFFRE DE THÈSE – Projet FlexMIEE

Modèles imbriqués et analyse de sensibilité pour le contrôle robuste des bioréacteurs à membranes à grande échelle

Projet ANR JCJC

Titre :	Modèles imbriqués et analyse de sensibilité pour le contrôle robuste des bioréacteurs à membranes à grande échelle.
Durée :	3 ans (2026–2029).
Encadrement :	Rachid OUARET (LGC – Équipe PSI).
Contact :	rachid.ouaret@toulouse-inp.fr

Laboratoire de Génie Chimique (LGC)

La thèse se déroulera à l'INP-ENSIACET, au sein du Laboratoire de Génie Chimique (LGC), unité mixte de recherche (UMR 5503) placée sous la triple tutelle du CNRS, de l'Institut National Polytechnique de Toulouse (Toulouse INP) et de l'Université Paul Sabatier (UPS). Le LGC est un laboratoire de référence en génie des procédés, rassemblant plus de 250 permanents et non-permanents répartis en plusieurs départements de recherche.

Le ou la doctorant(e) sera rattaché(e) au département Procédés et Systèmes Industriels (PSI), spécialisé dans la modélisation multi-échelle, l'optimisation et le contrôle des procédés. Le département développe des approches hybrides couplant modèles mécanistes, outils de statistiques avancées et intelligence artificielle, en étroite collaboration avec les acteurs industriels et les établissements publics d'assainissement. L'environnement de travail est fortement interdisciplinaire, à l'interface entre génie des procédés, data science, automatique et écologie industrielle.

Contexte scientifique

En France, plus de 22 000 stations d'épuration traitent annuellement environ 8,4 milliards de m³ d'eaux usées, mais moins de 1% de cette eau traitée est actuellement réutilisée. Dans un contexte de raréfaction croissante des ressources hydriques, le Plan Eau 2023 promeut activement la valorisation des eaux usées traitées [1]. Les bioréacteurs à membranes (BaM) se positionnent comme des procédés de traitement d'eaux usées reconnus pour leur excellente performance épuratoire et sont désormais déployés à grande échelle. Cependant, ces systèmes présentent un défi majeur : ils consomment au moins deux fois plus d'énergie que les procédés à boues activées conventionnels, principalement en raison de leurs besoins élevés en aération [3, 4]. Cette surconsommation énergétique devient particulièrement problématique dans le contexte de la crise énergétique actuelle, d'où l'urgence de développer des méthodologies permettant d'optimiser leur fonctionnement tout en respectant les contraintes environnementales strictes.

La complexité scientifique de cette problématique réside dans la nature intrinsèquement imbriquée des modèles MBR. Contrairement aux approches traditionnelles qui considèrent les systèmes comme des « monoblocs », les BaM à grande échelle impliquent une structure profondément imbriquée où la sortie d'un sous-modèle (biologique, physique ou énergétique) devient l'entrée d'un autre [2]. Cette architecture spécifique engendre des dynamiques multi-échelles et multi-fréquences complexes. Les phénomènes biologiques influencent directement la production de matières en suspension et d'EPS (Extracellular Polymeric Substances), qui conditionnent le colmatage des membranes, créant ainsi un système physico-biologique fortement couplé avec des rétroactions multiples. Cette structure imbriquée est essentielle pour capturer fidèlement le comportement réel des installations industrielles, mais elle complique considérablement l'analyse et la gestion des incertitudes.

Les incertitudes présentes dans ces systèmes sont multiples et dépendantes : elles portent sur l'influent (débit, DCO, NH_4^+ , MES), les conditions opératoires (paramètres de contrôle) et l'encrassement membranaire. Ces incertitudes évoluent dans le temps, sont décrites par des données hétérogènes souvent asynchrones ou incomplètes, et présentent des dépendances structurelles complexes qui ne peuvent pas être capturées par les méthodes statistiques classiques. En particulier, les dépendances entre variables (linéaires ou non linéaires, symétriques ou asymétriques) nécessitent des outils spécifiques tels que les copules pour une modélisation précise [5]. Cette complexité des incertitudes est exacerbée par la nature des données collectées : les stations d'épuration génèrent des séries temporelles à différentes échelles temporelles (de la seconde à la semaine), avec des pas de temps distincts entre les capteurs en ligne, les analyses de laboratoire et les historiques opératoires.

L'analyse de sensibilité représente un levier essentiel pour comprendre et optimiser ces systèmes, mais elle rencontre des défis majeurs dans le contexte des modèles imbriqués. Les méthodes traditionnelles d'analyse de sensibilité, développées pour des modèles monoblocs, ne permettent pas d'analyser finement les contributions des sous-modèles ni de quantifier correctement les sensibilités dans une structure imbriquée [6]. Une approche spécifique est donc nécessaire pour évaluer comment les variations des paramètres d'entrée affectent les sorties finales à travers l'ensemble de la chaîne modélisante. Cette analyse de sensibilité doit également intégrer la représentation symbolique des données (intervalles, histogrammes, distributions empiriques) pour capturer l'incertitude sans recourir à un lissage temporel qui perdrait de l'information critique [7].

Cœur de la thèse : Modélisation imbriquée, sensibilité et intelligence artificielle

Le cœur de cette thèse réside dans le développement d'une méthodologie novatrice articulant modélisation imbriquée, analyse de sensibilité et intelligence artificielle pour les bioréacteurs à membranes à grande échelle. Cette approche s'inscrit dans la continuité des travaux préliminaires menés au LGC sur la modélisation intégrée des MBR [8, 9] et s'appuie sur des résultats prometteurs obtenus récemment dans le domaine de l'analyse de sensibilité des modèles hybrides [10, 11].

La modélisation imbriquée constitue le fondement de cette thèse. Contrairement aux approches monoblocs traditionnelles, nous développerons une représentation graphique des interdépendances entre sous-modèles biologiques (ASM-SMP), physiques (RIS) et énergétiques, suivant une approche similaire à celle proposée par Touboul [12]. Cette structure imbriquée permettra de capturer fidèlement les rétroactions multiples entre les différents composants du système MBR, notamment entre les processus biologiques de dégradation des polluants et les phénomènes physiques de filtration membranaire. L'utilisation d'outils probabilistes tels que les copules [13] permettra de modéliser rigoureusement les dépendances stochastiques entre les variables influentes, tandis qu'une variante de l'indice d'interaction totale (TII) [14] sera mise en œuvre pour évaluer les dépendances stochastiques au sein du cadre imbriqué.

L'analyse de sensibilité occupe une place centrale dans cette thèse, avec pour objectif de développer des mesures de sensibilité spécifiquement adaptées aux structures imbriquées et multi-échelles. Nous exploiterons les outils classiques bien documentés dans la littérature spécialisée (approches basées sur les dérivées, distributions ou variogrammes) [15], tout en menant une réflexion approfondie sur la spécificité de la structure des modèles imbriqués pour adapter ces outils à notre contexte spécifique. L'effet Shapley pour l'analyse de sensibilité avec entrées dépendantes [16, 13] sera également considéré dans le cadre de ce projet, offrant une perspective innovante pour quantifier les contributions individuelles et conjointes des paramètres dans un contexte de dépendance.

L'intelligence artificielle jouera un rôle complémentaire et essentiel dans cette thèse, en particulier à travers le développement de réseaux de neurones informés par la physique (PINN) et de techniques d'interprétabilité des modèles d'apprentissage automatique. Cette approche hybride, qui combine les avantages des modèles mécanistes et des méthodes data-driven, a déjà fait ses preuves dans le domaine du génie des procédés [18] et spécifiquement pour le traitement des eaux usées [19]. Les récents travaux de Danesh et al. [17, 11] ont démontré l'efficacité des méthodes agnostiques de modèles (Accumulated Local Effects, Partial Dependence Plots) pour rendre interprétables les prédictions des réseaux de neurones, une exigence fondamentale pour l'adoption de ces outils par les opérateurs industriels.

L'apprentissage par renforcement (RL) constituera le troisième pilier de cette thèse, avec pour objectif de développer un cadre de contrôle en temps réel pour l'optimisation des opérations MBR. En s'appuyant sur les modèles imbriqués développés et les analyses de sensibilité réalisées, ce cadre RL ajustera dynamiquement les paramètres opérationnels en fonction de la variabilité du procédé. Des techniques d'estimation d'état, telles que le filtre de Kalman étendu (EKF), seront mises en œuvre pour améliorer la prise de décision RL en atténuant le bruit de mesure et en gérant les incertitudes du système [20, 21]. L'intégration du RL avec les modèles mécanistes garantira que la stratégie de contrôle reste explicable et applicable aux opérations réelles.

Objectifs scientifiques

Objectif 1 — Modèles MBR intégrés et imbriqués

L'objectif principal est de construire des modèles MBR intégrés et imbriqués qui couplent de manière cohérente les modèles biologiques, physiques et énergétiques, en s'appuyant à la fois sur des modèles mécanistes et des modèles de données. Un point central est la modélisation explicite des incertitudes dans chacun des sous-modèles, afin de développer une analyse de sensibilité réellement adaptée aux structures imbriquées et multi-échelles.

La nécessité de développement méthodologique est fondamentale dans ce contexte, car les méthodes traditionnelles d'analyse de sensibilité, conçues pour des modèles monoblocs, s'avèrent inadaptées aux structures imbriquées où la sortie d'un sous-modèle devient l'entrée d'un autre. Ce verrou scientifique nécessite le développement de mesures de sensibilité spécifiquement adaptées aux structures imbriquées, capables de quantifier non seulement l'impact direct des paramètres d'entrée sur la

sortie finale, mais aussi leur propagation à travers la chaîne modélisante complète.

Cette approche méthodologique innovante s’inscrit dans la lignée des travaux récents sur la modélisation intégrée des MBR [8, 9] et vise à combler le fossé entre les modèles mécanistes traditionnels et les approches purement data-driven. L’utilisation d’une représentation graphique des interdépendances, combinée à des outils probabilistes tels que les copules [13] et une variante de l’indice d’interaction totale (TII) [14], permettra de capturer rigoureusement les dépendances stochastiques au sein du cadre imbriqué. L’effet Shapley pour l’analyse de sensibilité avec entrées dépendantes [16, 13] sera également considéré comme une perspective innovante pour quantifier les contributions individuelles et conjointes des paramètres.

Objectif 2 — Contrôle avancé et hybride

Le second objectif concerne le contrôle avancé des BaM de grande échelle. Les approches classiques montrent leurs limites face aux incertitudes dynamiques et aux états non mesurables. Nous visons donc le développement de stratégies de commande prédictive robuste et adaptative, appuyées sur des observateurs avancés, et intégrant de manière raisonnée des approches hybrides modèles-données. L’apprentissage par renforcement est envisagé comme une brique adaptative complémentaire, et non comme un substitut aux modèles physiques. L’IA est ainsi utilisée de manière raisonnée, en complément des modèles mécanistes, pour garantir à la fois performance et interprétabilité. Cette approche hybride s’inspire des travaux récents sur le contrôle prédictif non linéaire des MBR [22] et des applications prometteuses de l’apprentissage par renforcement dans le traitement des eaux usées [23, 21].

Déroulé de la thèse

La thèse s’articule sur trois années avec une progression logique de la modélisation fondamentale à la validation industrielle.

Année 1 — Développement et validation d’un modèle MBR imbriqué intégrant les sous-modèles biologiques (ASM-SMP), physiques (RIS) et énergétiques. Mise en place d’une représentation graphique des interdépendances et réalisation d’une première analyse de sensibilité à l’aide d’outils classiques (dérivées, distributions). Développement initial de modèles hybrides PINN et application de méthodes d’interprétabilité (SHAP, ALE).

Année 2 — Approfondissement de l’analyse de sensibilité pour les structures imbriquées avec développement de mesures spécifiques (effets de Shapley pour variables dépendantes). Conception d’un cadre d’apprentissage par renforcement pour l’optimisation en temps réel des opérations MBR, intégrant des observateurs avancés (filtre de Kalman étendu) pour l’estimation d’états non mesurables.

Année 3 — Validation industrielle complète sur le cas d’étude SIAAP (station Seine Aval, 348 000 m³/jour). Consolidation de la méthodologie globale, rédaction de la thèse et préparation de publications dans des revues internationales à fort impact. Participation à des conférences scientifiques (SFGP, ESCAPE, ACODS) pour diffuser les résultats.

Cas d’étude industriel

La thèse bénéficiera d’un partenariat avec le Syndicat Interdépartemental pour l’Assainissement de l’Agglomération Parisienne (SIAAP), qui mettra à disposition des données réelles issues de la station d’épuration Seine Aval, d’une capacité de traitement de 348 000 m³/jour. Ce cas d’étude industriel de grande échelle offrira un terrain d’application idéal pour valider les méthodologies développées et démontrer leur pertinence opérationnelle. Les données comprennent des mesures en ligne, des

analyses de laboratoire et des historiques opératoires couvrant plusieurs années de fonctionnement, permettant une analyse approfondie des dynamiques multi-échelles et des incertitudes inhérentes aux procédés MBR industriels.

Profil recherché

- **Formation** : Diplôme de Master 2 ou équivalent (Bac+5) en mathématiques appliquées, statistique, génie des procédés, automatique, data science ou discipline connexe.
- **Compétences techniques** :
 - Probabilités et statistiques (analyse de sensibilité, processus stochastiques)
 - Modélisation de systèmes dynamiques et équations différentielles
 - Optimisation et contrôle optimal
 - Apprentissage automatique et apprentissage par renforcement
 - Programmation scientifique (Python, R, MATLAB, Julia)
- **Atouts** : Connaissances en génie des procédés ou traitement des eaux
- **Qualités personnelles** : Autonomie, rigueur scientifique, goût pour la recherche interdisciplinaire
- **Langues** : Anglais scientifique (lecture, écriture, communication orale)

Candidature

Dates de candidature : Les candidatures sont à envoyer à tout moment. Elles seront analysées au fil de l'eau et des entretiens seront proposés jusqu'au 30 avril 2026, jusqu'à ce que le poste soit pourvu.

Votre dossier doit comprendre :

- **Curriculum Vitae** détaillé (expériences, compétences techniques, projets)
- **Lettre de motivation** (maximum 1 page)
- **Relevés de notes** du Master (M1 et M2)
- **Rapport de stage** de Master 2 (si disponible)
- **Deux références académiques** (noms, fonctions, coordonnées)

Contact : rachid.ouaret@toulouse-inp.fr

Les entretiens auront lieu entre mars et avril 2026. La thèse débutera en septembre/octobre 2026.

Références

- [1] Ministère de la Transition écologique et de la Cohésion des territoires, « Plan Eau : le Gouvernement accélère la réutilisation des eaux usées traitées (REUT) », 30 août 2023.
- [2] Nadeem Kashif, « Complete mechanical model of a very large submerged membrane bioreactor », Institut National Polytechnique de Toulouse-INPT, 2022.
- [3] P. Krzeminski, J. H. van der Graaf, et J. B. van Lier, « Specific energy consumption of membrane bioreactor (MBR) for sewage treatment », *Water Sci. Technol.*, vol. 65, no 2, p. 380–392, 2012.
- [4] K. Tang et al., « The optimization and regulation of energy consumption for MBR process : A critical review », *J. Environ. Chem. Eng.*, vol. 10, no 5, p. 108406, 2022.
- [5] P. Zhang, « A novel feature selection method based on global sensitivity analysis with application in machine learning-based prediction model », *Appl. Soft Comput.*, vol. 85, p. 105859, 2019.
- [6] G. Mannina, T. F. Rebouças, A. Cosenza, et K. Chandran, « A plant-wide wastewater treatment plant model for carbon and energy footprint : Model application and scenario analysis », *J. Clean. Prod.*, vol. 217, p. 244–256, 2019.

- [7] R. Ouaret, P. Floquet, J.-P. Belaud, et S. Négny, « Interval-Valued Data Learning for Robustness of Energy Recovery Systems », in *Computer Aided Chemical Engineering*, vol. 50, Elsevier, 2021, p. 791–797.
- [8] G. Mannina et al., « Optimization of MBRs through integrated modelling : A state of the art », *J. Environ. Manage.*, vol. 370, p. 122720, 2024.
- [9] V. Sandoval-García et al., « Modeling MBR fouling : A critical review analysis towards establishing a framework for good modeling practices », *Water Res.*, vol. 268, p. 122611, 2025.
- [10] S. A. Halim, R. Ouaret, C. Albasi, et M. Alliet, « L’analyse de sensibilité par les réseaux de neurones artificiels (ANN) pour les bioréacteurs à membranes (BAM) », in *18ème congrès de la Société Française de Génie des Procédés*, 2022.
- [11] T. Danesh, R. Ouaret, P. Floquet, et S. Négny, « Hybridization of model-specific and model-agnostic methods for interpretability of Neural network predictions : Application to a power plant », *Comput. Chem. Eng.*, vol. 176, p. 108306, 2023.
- [12] A. Touboul, « Model of margin, margin sensitivity analysis and uncertainty quantification in graphs of functions in complex industrial systems », PhD Thesis, Marne-la-vallée, ENPC, 2021.
- [13] N. Benoumechiara et K. Elie-Dit-Cosaque, « Shapley effects for sensitivity analysis with dependent inputs : bootstrap and kriging-based algorithms », *ESAIM Proc. Surv.*, vol. 65, p. 266–293, 2019.
- [14] J. Fruth, O. Roustant, et S. Kuhnt, « Total interaction index : A variance-based sensitivity index for second-order interaction screening », *J. Stat. Plan. Inference*, vol. 147, p. 212–223, 2014.
- [15] S. Razavi et al., « The future of sensitivity analysis : an essential discipline for systems modeling and policy support », *Environ. Model. Softw.*, vol. 137, p. 104954, 2021.
- [16] A. B. Owen, « Sobol’ indices and Shapley value », *SIAM/ASA J. Uncertain. Quantif.*, vol. 2, no 1, p. 245–251, 2014.
- [17] T. Danesh, R. Ouaret, P. Floquet, et S. Négny, « Interpretability of neural networks predictions using Accumulated Local Effects as a model-agnostic method », in *Computer aided chemical engineering*, vol. 51, Elsevier, 2022, p. 1501–1506.
- [18] J. Mousa, S. Négny, R. Ouaret, A. Di Pretoro, et L. Montastruc, « Enhancing Unit Operation Design : Leveraging Neural Networks to Enforce Physical Hard Constraints », in *Computer Aided Chemical Engineering*, vol. 53, 2024, p. 547–552.
- [19] F. Guichard, M. Nadri-Wolf, R. Ouaret, et A. Azaïs, « Predicting sludge settleability in large wastewater treatment plants : a deep learning time series perspective », in *IWA 14th Specialized Conference on the Design, Operation and Economics of Large Wastewater Treatment Plants*, Budapest, 2024.
- [20] S. Syafiie, F. Tadeo, E. Martinez, et T. Alvarez, « Model-free control based on reinforcement learning for a wastewater treatment problem », *Appl. Soft Comput.*, vol. 11, no 1, p. 73–82, 2011.
- [21] K. Nam, S. Heo, J. Loy-Benitez, P. Ifaei, et C. Yoo, « An autonomous operational trajectory searching system for an economic and environmental membrane bioreactor plant using deep reinforcement learning », *Water Sci. Technol.*, vol. 81, no 8, p. 1578–1587, 2020.
- [22] X. Guo, P. Hong, et T.-M. Laleg-Kirati, « Nonlinear Model Predictive Control Design for BSM-MBR : Benchmark of Membrane Bioreactor », *IFAC-Pap.*, vol. 53, no 2, p. 16524–16530, 2020.
- [23] F. Hernández-del-Olmo, E. Gaudioso, R. Dormido, et N. Duro, « Energy and Environmental Efficiency for the N-Ammonia Removal Process in Wastewater Treatment Plants by Means of Reinforcement Learning », *Energies*, vol. 9, no 9, 2016.



PhD Position – Project FlexMIEE

Nested Modeling and Sensitivity Analysis for Robust Control of Large-Scale Membrane Bioreactors under Uncertainty ANR JCJC Project

Title: Nested Modeling and Sensitivity Analysis for Robust Control of
Large-Scale Membrane Bioreactors under Uncertainty.
Duration: 3 years (2026–2029).
Supervision: Rachid OUARET (LGC – PSI Team)
Contact: rachid.ouaret@toulouse-inp.fr

Laboratoire de Génie Chimique (LGC)

The PhD will be carried out at INP-ENSIACET, within the Laboratoire de Génie Chimique (LGC), a joint research unit (UMR 5503) under the joint supervision of CNRS, Institut National Polytechnique de Toulouse (Toulouse INP), and Université Paul Sabatier (UPS). LGC is a leading research laboratory in process engineering, hosting more than 250 permanent and non-permanent researchers across several departments.

The PhD candidate will be affiliated with the Procédés et Systèmes Industriels (PSI) department, which specializes in multi-scale modeling, process optimization, and control. The department develops hybrid approaches combining mechanistic models, advanced statistical tools, and artificial intelligence, in close collaboration with industrial partners and public wastewater management agencies. The working environment is highly interdisciplinary, bridging process engineering, data science, automatic control, and industrial ecology.

Scientific Context

In France, more than 22,000 wastewater treatment plants collect and treat approximately 8.4 billion m³ of wastewater annually, but less than 1% of this treated water is currently reused. In response to recent drought episodes and the increasing scarcity of water resources, the "Water plan (Plan Eau)" was introduced in 2023 to promote a more resilient and concerted management of water resources, including the valorization of treated wastewater reuse (REUT) [1]. Membrane bioreactors (MBRs) are recognized wastewater treatment processes known for their excellent purification performance and are now deployed in large-scale installations. However, these systems consume at least twice the energy of conventional activated sludge systems (CAS), primarily due to their high aeration requirements [3, 4]. This energy demand is further influenced by the concentration of Total Suspended Solids (TSS), which directly affects reactor volume and design. Addressing these issues requires advanced mathematical models capable of optimizing MBR performance while accounting for uncertainties in influent characteristics, operational conditions, effluent quality, and energy efficiency.

The scientific complexity of this issue lies in the intrinsically nested nature of MBR models. Unlike traditional approaches that consider systems as "single-block" entities, large-scale MBRs involve a deeply nested structure where the output of one sub-model (biological, physical, or energetic) becomes the input of another [2]. This specific architecture generates complex multi-scale and multi-frequency dynamics. Biological phenomena directly influence the production of suspended solids and EPS (Extracellular Polymeric Substances), which condition membrane fouling, creating a strongly coupled physico-biological system with multiple feedback loops. This nested structure is essential for faithfully capturing the real behavior of industrial installations, but it considerably complicates the analysis and management of uncertainties.

The uncertainties present in these systems are multiple and dependent: they concern the influent (flow rate, COD, NH₄⁺, TSS), operational conditions (control parameters), and membrane fouling. These uncertainties evolve over time, are described by heterogeneous data often asynchronous or incomplete, and exhibit complex structural dependencies that cannot be captured by classical statistical methods. In particular, dependencies between variables (linear or nonlinear, symmetric or asymmetric) require specific tools such as copulas for precise modeling [5]. This complexity of uncertainties is exacerbated by the nature of the collected data: wastewater treatment plants generate time series at different temporal scales (from seconds to weeks), with distinct time steps between online sensors, laboratory analyses, and operational histories.

Sensitivity analysis represents an essential lever for understanding and optimizing these systems, but it faces major challenges in the context of nested models. Traditional sensitivity analysis methods, developed for single-block models, do not allow for fine analysis of sub-model contributions nor correct quantification of sensitivities in a nested structure [6]. A specific approach is therefore necessary to evaluate how variations in input parameters affect final outputs through the entire modeling chain. This sensitivity analysis must also integrate symbolic data representation (intervals, histograms, empirical distributions) to capture uncertainty without resorting to temporal smoothing that would lose critical information [7].

Core of the PhD: Nested Modeling, Sensitivity and Artificial Intelligence

The heart of this PhD lies in the development of an innovative methodology articulating nested modeling, sensitivity analysis, and artificial intelligence for large-scale membrane bioreactors. This approach builds upon preliminary work conducted at LGC on integrated MBR modeling [8, 9] and leverages promising results recently obtained in the field of sensitivity analysis of hybrid models [10, 11].

Nested modeling constitutes the foundation of this PhD. Unlike traditional single-block approaches,

we will develop a graphical representation of interdependencies between biological (ASM-SMP), physical (RIS), and energetic sub-models, following an approach similar to that proposed by Touboul [12]. This nested structure will enable faithful capture of multiple feedback loops between different MBR system components, particularly between biological pollutant degradation processes and physical membrane filtration phenomena. The use of probabilistic tools such as copulas [13] will allow rigorous modeling of stochastic dependencies between influential variables, while a variant of the Total Interaction Index (TII) [14] will be implemented to assess stochastic dependencies within the nested framework.

Sensitivity analysis occupies a central position in this PhD, with the objective of developing sensitivity measures specifically adapted to nested and multi-scale structures. We will deploy classical tools well-documented in specialized literature (derivative-based, distribution-based, or variogram-based approaches) [15], while conducting an in-depth reflection on the specificity of nested model structures to tailor these tools to our specific context. The Shapley effect for sensitivity analysis with dependent inputs [16, 13] will also be considered within the framework of this project, offering an innovative perspective for quantifying individual and joint contributions of parameters in a dependency context.

Artificial intelligence will play a complementary and essential role in this PhD, particularly through the development of Physics-Informed Neural Networks (PINN) and machine learning model interpretability techniques. This hybrid approach, which combines the advantages of mechanistic models and data-driven methods, has already proven its worth in the field of process engineering [18] and specifically for wastewater treatment [19]. Recent work by Danesh et al. [17, 11] has demonstrated the effectiveness of model-agnostic methods (Accumulated Local Effects, Partial Dependence Plots) for making neural network predictions interpretable, a fundamental requirement for the adoption of these tools by industrial operators.

Reinforcement learning (RL) will constitute the third pillar of this PhD, with the objective of developing a real-time control framework for optimizing MBR operations. Building upon the nested models developed and sensitivity analyses performed, this RL framework will dynamically adjust operational parameters based on process variability. State estimation techniques, such as the Extended Kalman Filter (EKF), will be implemented to enhance RL decision-making by mitigating measurement noise and handling system uncertainties [20, 21]. The integration of RL with mechanistic models will ensure that the control strategy remains explainable and applicable to real-world operations.

Scientific Objectives

Objective 1 — Integrated and Nested MBR Models

The main objective is to build integrated and nested MBR models that coherently couple biological, physical, and energetic models, relying on both mechanistic models and data-driven models. A central point is the explicit modeling of uncertainties in each of the sub-models, in order to develop a sensitivity analysis truly adapted to nested and multi-scale structures.

The necessity for methodological development is fundamental in this context, as traditional sensitivity analysis methods, designed for single-block models, prove inadequate for nested structures where the output of one sub-model becomes the input of another. This scientific barrier requires the development of sensitivity measures specifically adapted to nested structures, capable of quantifying not only the direct impact of input parameters on the final output, but also their propagation throughout the complete modeling chain.

This innovative methodological approach builds upon recent work on integrated MBR modeling [8, 9] and aims to bridge the gap between traditional mechanistic models and purely data-driven approaches. The use of a graphical representation of interdependencies, combined with probabilistic

tools such as copulas [13] and a variant of the Total Interaction Index (TII) [14], will enable rigorous capture of stochastic dependencies within the nested framework. The Shapley effect for sensitivity analysis with dependent inputs [16, 13] will also be considered as an innovative perspective for quantifying individual and joint contributions of parameters in a dependency context.

Objective 2 — Advanced and Hybrid Control

The second objective concerns advanced control of large-scale MBRs. Classical approaches show their limitations in the face of dynamic uncertainties and unmeasurable states. We therefore aim to develop robust and adaptive predictive control strategies, supported by advanced observers, and integrating hybrid model-data approaches in a reasoned manner. Reinforcement learning is considered as a complementary adaptive component, not as a substitute for physical models. AI is thus used in a reasoned way, complementing mechanistic models, to guarantee both performance and interpretability. This hybrid approach draws inspiration from recent work on nonlinear predictive control of MBRs [22] and promising applications of reinforcement learning in wastewater treatment [23, 21].

PhD Timeline

The PhD unfolds over three years with a logical progression from fundamental modeling to industrial validation.

Year 1 — Development and validation of an integrated nested MBR model coupling biological (ASM-SMP), physical (RIS), and energetic sub-models. Implementation of a graphical representation of interdependencies and preliminary sensitivity analysis using classical tools (derivatives, distributions). Initial development of hybrid Physics-Informed Neural Networks (PINN) and application of interpretability methods (SHAP, ALE).

Year 2 — Deepening of sensitivity analysis for nested structures through the development of specific measures (Shapley effects for dependent variables). Design of a reinforcement learning framework for real-time optimization of MBR operations, integrating advanced state observers (Extended Kalman Filter) for unmeasurable state estimation.

Year 3 — Full industrial validation on the SIAAP case study (Seine Aval plant, 348,000 m³/day). Consolidation of the global methodology, thesis writing, and preparation of publications in high-impact international journals. Participation in scientific conferences (SFGP, ESCAPE, ACODS) to disseminate results.

Industrial Case Study

The PhD will benefit from a privileged partnership with the Syndicat Interdépartemental pour l'Assainissement de l'Agglomération Parisienne (SIAAP), which will provide real data from the Seine Aval wastewater treatment plant, with a treatment capacity of 348,000 m³/day. This large-scale industrial case study will offer an ideal application ground for validating the methodologies developed and demonstrating their operational relevance. The data include online measurements, laboratory analyses, and operational histories covering several years of operation, enabling in-depth analysis of multi-scale dynamics and inherent uncertainties in industrial MBR processes.

Candidate Profile

- **Education:** Master's degree or equivalent (5 years) in applied mathematics, statistics, process engineering, control engineering, data science, or related field.

- **Technical skills:**
 - Probability and statistics (sensitivity analysis, stochastic processes)
 - Dynamical systems modeling and differential equations
 - Optimization and optimal control
 - Machine learning and reinforcement learning
 - Scientific programming (Python, R, MATLAB, Julia)
- **Assets:** Knowledge in process engineering or water treatment would be a plus
- **Personal qualities:** Autonomy, scientific rigor, taste for interdisciplinary research
- **Languages:** Scientific English (reading, writing, oral communication)

Application Procedure

Application dates: Applications can be submitted at any time. They will be reviewed on a rolling basis, and interviews will be offered until April 30, 2026, until the position is filled.

Your application must include:

- **Detailed Curriculum Vitae** (experiences, technical skills, projects)
- **Motivation letter** (maximum 1 page)
- **Master's degree transcripts** (M1 and M2)
- **Master's thesis report** (if available)
- **Two academic references** (names, positions, contact details)

Contact: rachid.ouaret@toulouse-inp.fr

Interviews will take place between March and April 2026. The PhD will start in September/October 2026.

References

- [1] Ministère de la Transition écologique et de la Cohésion des territoires, « Plan Eau: le Gouvernement accélère la réutilisation des eaux usées traitées (REUT) », 30 août 2023.
- [2] Nadeem Kashif, « Complete mechanical model of a very large submerged membrane bioreactor », Institut National Polytechnique de Toulouse-INPT, 2022.
- [3] P. Krzeminski, J. H. van der Graaf, et J. B. van Lier, « Specific energy consumption of membrane bioreactor (MBR) for sewage treatment », *Water Sci. Technol.*, vol. 65, no 2, p. 380–392, 2012.
- [4] K. Tang et al., « The optimization and regulation of energy consumption for MBR process: A critical review », *J. Environ. Chem. Eng.*, vol. 10, no 5, p. 108406, 2022.
- [5] P. Zhang, « A novel feature selection method based on global sensitivity analysis with application in machine learning-based prediction model », *Appl. Soft Comput.*, vol. 85, p. 105859, 2019.
- [6] G. Mannina, T. F. Rebouças, A. Cosenza, et K. Chandran, « A plant-wide wastewater treatment plant model for carbon and energy footprint: Model application and scenario analysis », *J. Clean. Prod.*, vol. 217, p. 244–256, 2019.
- [7] R. Ouaret, P. Floquet, J.-P. Belaud, et S. Négny, « Interval-Valued Data Learning for Robustness of Energy Recovery Systems », in *Computer Aided Chemical Engineering*, vol. 50, Elsevier, 2021, p. 791–797.
- [8] G. Mannina et al., « Optimization of MBRs through integrated modelling: A state of the art », *J. Environ. Manage.*, vol. 370, p. 122720, 2024.
- [9] V. Sandoval-García et al., « Modeling MBR fouling: A critical review analysis towards establishing a framework for good modeling practices », *Water Res.*, vol. 268, p. 122611, 2025.
- [10] S. A. Halim, R. Ouaret, C. Albasi, et M. Alliet, « L'analyse de sensibilité par les réseaux de neurones artificiels (ANN) pour les bioréacteurs à membranes (BAM) », in *18ème congrès de la Société Française de Génie des Procédés*, 2022.

- [11] T. Danesh, R. Ouaret, P. Floquet, et S. Négny, « Hybridization of model-specific and model-agnostic methods for interpretability of Neural network predictions: Application to a power plant », *Comput. Chem. Eng.*, vol. 176, p. 108306, 2023.
- [12] A. Touboul, « Model of margin, margin sensitivity analysis and uncertainty quantification in graphs of functions in complex industrial systems », PhD Thesis, Marne-la-vallée, ENPC, 2021.
- [13] N. Benoumechiara et K. Elie-Dit-Cosaque, « Shapley effects for sensitivity analysis with dependent inputs: bootstrap and kriging-based algorithms », *ESAIM Proc. Surv.*, vol. 65, p. 266–293, 2019.
- [14] J. Fruth, O. Roustant, et S. Kuhnt, « Total interaction index: A variance-based sensitivity index for second-order interaction screening », *J. Stat. Plan. Inference*, vol. 147, p. 212–223, 2014.
- [15] S. Razavi et al., « The future of sensitivity analysis: an essential discipline for systems modeling and policy support », *Environ. Model. Softw.*, vol. 137, p. 104954, 2021.
- [16] A. B. Owen, « Sobol’ indices and Shapley value », *SIAM/ASA J. Uncertain. Quantif.*, vol. 2, no 1, p. 245–251, 2014.
- [17] T. Danesh, R. Ouaret, P. Floquet, et S. Negny, « Interpretability of neural networks predictions using Accumulated Local Effects as a model-agnostic method », in *Computer aided chemical engineering*, vol. 51, Elsevier, 2022, p. 1501–1506.
- [18] J. Mousa, S. Negny, R. Ouaret, A. Di Pretoro, et L. Montastruc, « Enhancing Unit Operation Design: Leveraging Neural Networks to Enforce Physical Hard Constraints », in *Computer Aided Chemical Engineering*, vol. 53, 2024, p. 547–552.
- [19] F. Guichard, M. Nadri-Wolf, R. Ouaret, et A. Azaïs, « Predicting sludge settleability in large wastewater treatment plants: a deep learning time series perspective », in *IWA 14th Specialized Conference on the Design, Operation and Economics of Large Wastewater Treatment Plants*, Budapest, 2024.
- [20] S. Syafie, F. Tadeo, E. Martinez, et T. Alvarez, « Model-free control based on reinforcement learning for a wastewater treatment problem », *Appl. Soft Comput.*, vol. 11, no 1, p. 73–82, 2011.
- [21] K. Nam, S. Heo, J. Loy-Benitez, P. Ifaei, et C. Yoo, « An autonomous operational trajectory searching system for an economic and environmental membrane bioreactor plant using deep reinforcement learning », *Water Sci. Technol.*, vol. 81, no 8, p. 1578–1587, 2020.
- [22] X. Guo, P. Hong, et T.-M. Laleg-Kirati, « Nonlinear Model Predictive Control Design for BSM-MBR: Benchmark of Membrane Bioreactor », *IFAC-Pap.*, vol. 53, no 2, p. 16524–16530, 2020.
- [23] F. Hernández-del-Olmo, E. Gaudioso, R. Dormido, et N. Duro, « Energy and Environmental Efficiency for the N-Ammonia Removal Process in Wastewater Treatment Plants by Means of Reinforcement Learning », *Energies*, vol. 9, no 9, 2016.