

Offres TFE 2023-2024

Purecontrol

Présentation de l'entreprise

Purecontrol, entreprise fondée en 2017, a pour vocation de **réduire l'impact énergétique** des installations industrielles - ainsi que leur coût de fonctionnement - par une **optimisation** du **pilotage** de leurs équipements. Afin de pouvoir proposer à un maximum d'industriels des solutions de pilotage avancées, il est important de développer des approches minimisant le plus possible les coûts de développement et de déploiement. Les avancées récentes en *machine learning* offrent de nouveaux outils permettant cette réduction des coûts. Purecontrol rentre dans cette tendance en utilisant au maximum ces outils dans une approche de pilotage **basée données**. Cette approche basée données se traduit par la modélisation de systèmes dynamiques par *machine learning* ainsi que la synthèse de contrôleurs prédictifs tels que l'apprentissage par renforcement [7], la programmation dynamique ou les *MPC* [5] à partir des modèles appris. [1, 9]

Aujourd'hui, la solution Purecontrol est majoritairement déployée chez des acteurs du milieu de l'eau : optimisation de l'aération dans les stations d'épuration, ou du pompage et de la distribution d'eau potable. D'autres projets industriels sont également en cours de déploiement sur des sites de galvanisation, des incinérateurs, des systèmes de chambres froides, etc.

Au sein d'une entreprise très dynamique, jeune et ambitieuse, vous pourrez mettre en pratique les acquis de votre formation sur des sujets vertueux et innovants.

Sujets proposés

Prédiction de la concentration en phosphore en station d'épuration

Outre le traitement des composés azotés (déjà adressé par le pilotage intelligent Purecontrol), l'une des missions des stations d'épuration est l'élimination du phosphore des eaux usées. Il existe pour ce faire différentes techniques plus ou moins coûteuses [2], et les mécanismes de certaines d'entre elles sont encore assez mal comprises par les experts. [6]

L'objectif du TFE sera de modéliser la dynamique de l'élimination du phosphore dans les bassins des STEP (STations d'EPuration) en utilisant des modèles basés données, prédisant donc l'évolution du phosphore en fonction des équipements pilotables par Purecontrol et de différentes variables exogènes à définir. On travaillera dans un premier temps sur les quelques sites équipés de capteurs de phosphore. Ces derniers étant coûteux à installer et à entretenir, l'immense majorité des STEP n'en dispose pas ; dans ce cas, seules quelques mesures ponctuelles sont réalisées en sortie d'ouvrage. Il s'agira donc dans un second temps de trouver une façon de prédire ces relevés ponctuels. Cette tâche étant particulièrement difficile du fait de la parcimonie des données à disposition, on essaiera de capitaliser sur les sites équipés étudiés en première partie, et on pourra éventuellement s'aider de connaissances physico-chimiques sur le sujet. On n'oubliera pas tout au long du projet que ces modèles doivent servir à l'élaboration d'une stratégie de contrôle optimal, permettant de respecter différentes contraintes métier et légales (portant sur un seuil maximal de phosphore en sortie de d'ouvrage), tout en minimisant les coûts de fonctionnement de la station.

Développement d'algorithme(s) de détection d'anomalies

Purecontrol collecte auprès de chacun de ses clients des données temporelles, issues de milliers d'automates et de capteurs, qui servent après quelques mois de collecte à entraîner des modèles *ad hoc* afin de piloter intelligemment les systèmes concernés. Il arrive parfois que le fonctionnement dudit système sorte de la normalité (relative à un modèle sous-jacent, lui-même à définir) : panne d'un aérateur, blocage d'une conduite, dérive d'une sonde, etc. Il est nécessaire de pouvoir détecter automatiquement de telles anomalies pour plusieurs raisons. En premier lieu, si un dysfonctionnement a été détecté, il s'agira de remonter l'information aux personnes adéquates (exploitant de l'usine, par exemple), et éventuellement d'arrêter ou de modifier la façon dont sont pilotés ses équipements. Au

contraire, si cette anomalie correspond simplement à un *data drift* (dérive ou changement de sonde, modification de la distribution des données induit par une nouvelle façon de réguler, etc.), on pourra déclencher un ré-entraînement des modèles utilisés pour le pilotage. On pourra aussi, s'il s'agit d'une anomalie ponctuelle que l'on peut assimiler à une période aberrante, retirer automatiquement ce laps de temps du jeu de données d'entraînement des futurs modèles.

L'objectif du TFE sera, tout d'abord, de réaliser un état de l'art des méthodes de détection d'anomalies existantes. [4] Par la suite, il s'agira de développer un ou plusieurs algorithmes de détection d'anomalies répondant aux besoins précédemment évoqués. Une attention toute particulière sera portée à la généralité des techniques utilisées, ainsi qu'à leur robustesse ; en effet, dans un souci de scalabilité, cet outil devra être fonctionnel sur un maximum de clients possibles et ce, sans qu'un *fine tuning* laborieux ne soit nécessaire pour chaque site. En outre, pour qu'on y voit une réelle utilité, il faudra veiller à ce que le nombre de faux positifs soit minimal.

Modélisation de l'impact de phénomènes environnementaux sur les débits d'eau

Actuellement, une très grande partie des clients de Purecontrol appartient au monde du traitement de l'eau. L'eau usée en zone urbaine, notamment, circule dans des réseaux afin d'être acheminée vers des STations d'EPuration (STEP) ou des Usines d'Eau Potable (UEP). Le débit d'eau entrant est un paramètre particulièrement important à prendre en compte dans la façon de piloter un ouvrage de traitement d'eau.

Afin de s'extraire de la complexité des modèles experts de prédiction de débit [3], on souhaite utiliser une approche basée données. Lorsqu'il s'agit d'une STEP urbaine, il s'avère que l'on peut prédire de manière à peu près correcte le débit entrant à l'aide de modèles de séries temporelles [8] : on observe généralement des *patterns* très réguliers, l'utilisation de l'eau par les habitants étant régie par un certain nombre d'habitudes. Cependant, deux autres facteurs non périodiques ont un impact sur le débit entrant : la pluie et les infiltrations d'eau provenant de nappes phréatiques.

Le TFE aura pour objectif d'imaginer une façon d'utiliser ces variables exogènes afin d'améliorer les performances des modèles de prédiction de débits existants. En plus de donner de meilleures prédictions et par conséquent d'améliorer les performances de pilotage, cette nouvelle modélisation permettra d'identifier les portions de réseau où l'infiltration est trop importante (et par la suite d'effectuer les réparations nécessaires), évitant ainsi des explorations du réseau longues et très coûteuses. La difficulté de ce projet résidera notamment dans le fait que la modélisation devra s'adapter automatiquement à toutes les typologies de STEP urbaines existantes - certaines étant impactées par la pluie instantanément, d'autres dont l'impact de la pluie est observé plusieurs jours après l'épisode pluvieux, etc.

Travail à effectuer et profil recherché

Pour les différents sujets proposés, le travail à effectuer consistera en :

- la montée en compétence sur les méthodes en lien avec le sujet de stage,
- la lecture d'**articles scientifiques**,
- la **formulation mathématique** de la solution proposée,
- l'**implémentation** et validation de la solution proposée.

Les compétences recherchées chez les candidat-e-s sont les suivantes :

- maîtrise des **outils mathématiques liés au *machine learning***,
- connaissance de Python et de ses bibliothèques de *machine learning* (scikit-learn, TensorFlow ou PyTorch, etc.),
- si possible, quelques notions en SQL et InfluxDB,
- autonomie, notamment concernant l'exploration de la littérature scientifique,
- ainsi qu'une sensibilité aux **enjeux environnementaux** et climatiques !

Contact et modalités

Le TFE se déroulera dans les locaux de l'antenne liégeoise de Purecontrol, à la Grand Poste, et sera encadré par l'équipe R&D. Pour soumettre votre candidature, vous pouvez envoyer un mail à : noemie.godbillot@purecontrol.com.

Références

- [1] Pouria Bahramnia, Seyyed Mohammad Hosseini Rostami, Jin Wang, and Gwang jun Kim. Modeling and controlling of temperature and humidity in building heating, ventilating, and air conditioning system using model predictive control. *Energies*, 12(24), dec 2019.
- [2] Gaëlle Deronzier and Jean-Marc Choubert. Traitement du phosphore dans les petites stations d’épuration à boues activées. *Fonds National pour le Développement des Adductions d’Eau potable*, 29, 2004.
- [3] Krist V. Gernaey, Xavier Flores-Alsina, Christian Rosen, Lorenzo Benedetti, and Ulf Jeppsson. Dynamic influent pollutant disturbance scenario generation using a phenomenological modelling approach. *Environmental Modelling & Software*, 26(11) :1255–1267, 2011.
- [4] Songqiao Han, Xiyang Hu, Hailiang Huang, Mingqi Jiang, and Yue Zhao. Adbench : Anomaly detection benchmark, 2022.
- [5] J.B. Rawlings, D.Q. Mayne, and M. Diehl. *Model Predictive Control : Theory, Computation, and Design*. Nob Hill Publishing, 2017.
- [6] Robert J. Seviour, Takashi Mino, and Motoharu Onuki. The microbiology of biological phosphorus removal in activated sludge systems. *FEMS Microbiology Reviews*, 27(1) :99–127, 04 2003.
- [7] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning : An Introduction*. The MIT Press, second edition, 2018.
- [8] Sean J Taylor and Benjamin Letham. Forecasting at scale. *The American Statistician*, 72(1) :37–45, 2018.
- [9] Tianshu Wei, Yanzhi Wang, and Qi Zhu. Deep Reinforcement Learning for Building HVAC Control. *Proceedings - Design Automation Conference*, Part 12828, 2017.