Projet Wathoh-Wathoh

Outil d'aide à l'exploitation des STEP industrielles

Synthèse

SPEC ENVIRONNEMENT
26, rue Jean Mermoz
31 140 FONBEAUZARD
Tél : 05.34.66.61.10
Fax : 05.61.46.71.48
www.spec-environnement.fr

Date : 14/02/2024
Réf : S24WW01V0

Table des matières

1. Préambule 4
   a. Architecture du rapport 4
   b. Remerciements 4
2. Le projet 5
   a. Contexte du projet 5
   b. Objectif du projet 5
   c. Résultats attendus 6
   d. Déroulement du projet 7
      d-1. Partenaires 7
      d-2. Comité de pilotages 7
      d-3. Planning 8
3. L'application 9
   a. Architecture 9
   b. Connexion et navigation 11
   c. Vue Traitement 12
   d. Vue Résultats 13
   e. Vue Mesures 14
   f. Vue Guide 15
      f-1. A faire 15
      f-2. Analyses process 15
      f-3. Prévisions 16
      f-4. Conseils réglages 16
4. Modélisation numérique des unités de traitements 17
   a. Pilotes de traitement 17
      a-1. Construction 18
      a-2. Suivi des pilotes 22
      a-3. Bilan 22
   b. L'Intelligence Artificielle 23
      b-1. Preuve du concept (PoC) 23
      b-2. Validation avec les pilotes 24
      b-3. Choix des outils de modélisation 24
      b-4. Mise en production 25
5. Le déploiement 26

a. Préambule 26
b. FORESA 26
c. Industriel n°2 27
d. Industriel n°3 28
e. Industriel n°4 29
f. Industriel n°5 29
g. Enseignements des déploiements réalisés 30
g-1. Typologie de traitement 30
g-2. Quantité et qualité des données 30
g-3. Durée nécessaire au déploiement de l'IA 30
g-4. Wathoh-Wathoh = outil de terrain 31
6. Synthèse du projet 32

1. Préambule

a. *Architecture du rapport*

Nous avons pris le parti de présenter, dans un premier temps, le produit finalisé développé dans le cadre du projet : l'application Wathoh-Wathoh. Cette première partie constitue, en soi, un véritable manuel d'utilisateur qui sera déployé avec l'application. Puis dans un second temps, nous présentons les moyens mis en œuvre dans ce projet, avec un focus sur le développement de l'algorithme d'intelligence artificielle, et sur le déploiement de l'application sur différents sites industriels.

Ainsi, après un rappel du contexte (chapitre 1), l'application est présentée (chapitre 2), puis les moyens mis en œuvre sont détaillés (chapitres 3 et 4).

b. *Remerciements*

Nous tenons à remercier ceux qui, au sein de l'Agence de l'Eau, ont cru en notre projet dès le départ, et en ont assuré la présentation et la défense auprès de la Commission d'aide : Lucile Grémy, Ariette Sourzac, et Eric Gouzènes. Il est évident que les aides financières apportées ont clairement permis d'aller au bout de ce développement.

Nous remercions également Jean-Marc Pasanau et Arnaud Planche pour leur implication et leur enthousiasme dans ce projet, et pour nous avoir supporté toutes les semaines, avec nos bidons d'effluents !

2. Le projet

a. *Contexte du projet*

Le traitement des eaux usées industrielles est par nature complexe. En effet, chaque site industriel est singulier en raison :

- De son implantation,
- De son historique,
- Des procédés de fabrication et de production qui lui sont propres,
- Du contexte réglementaire,
- Des perspectives d'évolution,
- Du milieu environnant (naturel, sociétal, urbain…),
- Des équipements et infrastructures dont il dispose,
- De la stratégie environnementale menée,
- Etc…

La conséquence de cette singularité contextuelle est une singularité des effluents générés, qui nécessitent d'être traités avant rejet. Chaque couple « effluent + moyen de traitement » est unique, et le retour d'expérience pour la conception et l'exploitation est souvent insuffisant pour que le traitement en place soit optimal.

La conception et l'exploitation des outils de traitement de ces effluents requièrent :

- Des compétences multidisciplinaires,
- L'intégration, la capitalisation, la centralisation et l'analyse de nombreuses données,
- La prise en compte de marges de sécurité sur le dimensionnement pour pallier les « inconnues » de design, se traduisant par des coûts d'investissement majorés,
- Des moyens (humains et matériels) de contrôle du process épuratoire importants, impliquants là aussi des surcoûts d'exploitation.

Les besoins des industriels, dans le domaine du traitement des effluents, sont donc d'augmenter en continu leurs connaissances et compétences en la matière, **pour maîtriser leur impact sur le milieu environnant** tout en maîtrisant leurs dépenses, et **en réduisant les risques de pollution** du milieu et de non-conformités réglementaires.

b. *Objectif du projet*

Dans ce cadre, le projet « WATHOH-WATHOH » a pour objectif de développer un outil d'assistance à l'exploitation des stations d'épuration industrielles sous la forme d'une application pour smartphone et tablette.

A partir d'informations collectées sur site (capteurs connectés, analyses réalisées, système d'automatisme existant…) et de « l'intelligence process» acquise par le système, l'application guide les

exploitants dans les actions à réaliser pour améliorer la qualité des rejets, réduire l'empreinte environnementale globale (meilleure gestion des réactifs et de l'énergie) et prévenir les incidents.

Le moteur d'intelligence process s'appuie sur de l'IA et sur les compétences d'experts métiers.

L'application, adaptée au cas par cas, permet :

- De collecter les informations (capteurs connectés (IoT), et saisie directe des paramètres),
- De sauvegarder les informations sur une base de données via un cloud dédié,
- De visualiser le process épuratoire et les paramètres de fonctionnement,
- D'analyser en continu l'ensemble de ces données,
- De restituer les paramètres de pilotage en les hiérarchisant par priorité.

L'analyse des données collectées s'appuie sur :

- L'intelligence artificielle pour traiter une masse importante de données et proposer des corrélations possibles,
- Une expertise process d'ingénierie du traitement des eaux propre aux effluents industriels.

C. *Résultats attendus*

L'application doit permettre une visualisation aisée :

- Des données en temps réel,
- Des courbes de résultats, et des valeurs potentiellement critiques,
- Des corrélations possibles entre différents paramètres,
- Des actions à réaliser, par ordre de priorité, pour optimiser le fonctionnement ou corriger les anomalies.

d. *Déroulement du projet*

d-1. Partenaires

Les partenaires du projet sont les suivants :

- SPEC Environnement est l'entreprise principale portant le projet. Elle est un bureau d'étude reconnu et spécialisé dans le traitement des eaux industrielles. Elle pilote la réalisation du projet et assure l'exploitation des pilotes de traitement. Elle préside le comité de pilotage.

- L'Agence de l'eau Adour-Garonne subventionne le projet et participe au comité de pilotage.

- FORESA est un industriel de la chimie produisant principalement des colles formol-urée. Il dispose d'une station d'épuration très récente, dimensionnée principalement pour le traitement de la pollution azotée (station d'épuration biologique de type nitrification/dénitrification), exploitée par un prestataire extérieur. Les données de fonctionnement sont enregistrées sur différents supports (automate, supervision, tableurs Excel, cahier d'exploitation…) mais non corrélées et analysées. Il met à disposition son site, et participe financièrement au déploiement de l'outil à hauteur de 20 k€.

- CUBHOH est l'entreprise en charge du développement de l'application.

d-2. Comité de pilotages

Pour favoriser le bon déroulement du Projet, un Comité de pilotage a été créé. Il est constitué de l'Agence de l'eau Adour Garonne, de FORESA France, de SPEC ENVIRONNEMENT et de CUBHOH. Le comité est présidé par SPEC ENVIRONNEMENT.

Le Comité de pilotage a pour but :

- De suivre le bon déroulement du Projet, et de valider les jalons au fur et à mesure,
- De s'assurer que les conditions sont remplies pour l'exécution du Projet,
- Proposer et décider de mettre en œuvre des actions communes visant à lever les difficultés rencontrées lors de l'exécution du Projet,
- Décider ensemble de la communication qui sera réalisée au sujet du Projet, et de l'utilisation des noms des Partenaires.

Le comité s'est réuni lors de chaque étape principale du projet.

d-3. Planning

Le projet s'est déroulé entre février 2020 et décembre 2023. Il s'est décomposé en deux principales phases :

- Phase 1 : construction de l'application
- Phase 2 : déploiement de l'outil

❖ *Phase 1 : construction de l'application*

L'objectif de cette phase a été de développer le « squelette » de l'application et de construire et valider les modèles de prédiction de fonctionnement des unités de traitement.

Pour cela, nous avons construit 3 pilotes de traitement :

- Un pilote reproduisant une station d'épuration d'effluents agroalimentaires basique (industrie viticole), de type méthanisation + boues activées pour le traitement de la pollution organique, construit en impression 3D et équipé de différents capteurs connectés. L'alimentation était composée d'effluents synthétisés in situ pour cadrer les variations de qualité.
- Deux pilotes jumeaux miniatures de la station d'épuration du site de FORESA France : un pilote servant d'étalon pour caler la reproductibilité de la station réelle et un autre permettant de simuler et d'entraîner des variations de fonctionnement (alimentation, paramétrage, incidents…). L'alimentation était réalisée par l'amenée des effluents de FORESA 1 à 2 fois par semaine par nos soins (réaliste par la capacité du bassin tampon de FORESA qui représente entre 10 et 15 jours de temps de séjour).

Ces 3 pilotes ont été construits et suivis dans les locaux de SPEC Environnement.

Un suivi journalier des analyses et des paramètres de fonctionnement a été réalisé et capitalisé grâce à un système informatique connecté de contrôle / commande. Plusieurs configurations de charges, d'incidents, de conditions opératoires ont été testées pour « entraîner » les modèles.

Le suivi et l'exploitation de ces pilotes nous ont aiguillé dans la conception de l'application et dans l'élaboration des fonctionnalités nécessaires aux exploitants utilisateurs.

A partir des données historiques enregistrées depuis 2 ans sur l'usine de FORESA et de celles enregistrées sur les pilotes durant les 12 mois de suivi, nous avons pu construire et valider les modèles de prédiction.

❖ *Phase 2 : déploiement de l'outil*

Afin de valider la pertinence de l'application et des modèles créés, nous avons déployé l'outil sur 5 sites industriels, dont FORESA.

Chaque site nous a permis de valider, corriger et améliorer différents aspects et fonctionnalités de notre outil.

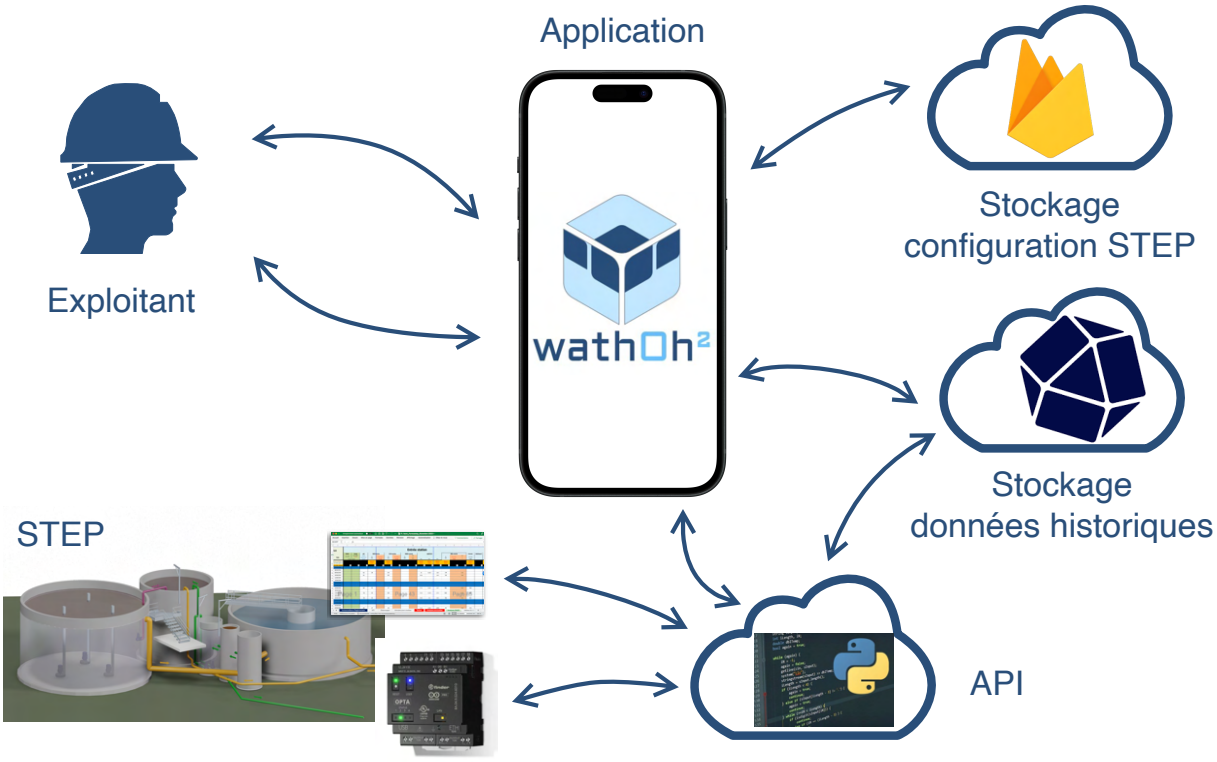
3. L'application

Disponible sur téléphone et tablette, elle permet les fonctionnalités principales suivantes :

- Visualisation du process de traitement avec les données collectées,
- Visualisation des principaux indicateurs de fonctionnement,
- Saisie ergonomique des valeurs de suivi journalier de la STEP (y compris scan des analyses lues et calcul automatisé pour certaines mesures (MES, dilutions, …)),
- Reporting de l'historique de tous les paramètres (et la possibilité de combiner les courbes ensemble)
- Fourniture de conseils hiérarchisés (issus de l'analyse d'IA ou de l'expertise process)
- Prédictions des valeurs de sortie du traitement par l'IA à 6 ou 7 jours,
- Intégration en automatique des données (supervision, fichiers Excel).

a. *Architecture*

Le schéma suivant présente l'architecture globale de notre outil.



Les rôles et fonctions de chaque élément sont détaillés ci-dessous.

❖ *L'application*



Elle permet à l'exploitant de visualiser le fonctionnement de la STEP, d'obtenir les prédictions sur les paramètres de sortie et les conseils de réglage.

L'exploitant reçoit des notifications sur des alertes déclenchées ou sur de nouveaux conseils de réglages.

L'exploitant peut également saisir directement les valeurs relevées sur sites, des analyses ou des observations.

❖ *Le stockage des configurations STEP*



Afin de prévoir une application modulaire et adaptative, nous avons décomposé chaque station d'épuration en divers composants (ouvrages, capteurs, indicateurs …). Chaque type de composant possède une liste commune d'attributs. La structure de l'application est donc identique pour toutes les stations, seuls les composants sont à paramétrer. Il est donc aisé d'ajouter des éléments (un nouveau capteur ou un nouvel indicateur) ou changer certains paramétrages (seuils de rejet, …) à la demande sans devoir modifier le programme de l'application.

Cette configuration est stockée sur le Cloud. Elle est téléchargée à chaque ouverture de l'application.

❖ *Le stockage des données historiques*



L'outil permet de saisir, enregistrer, visualiser et utiliser les données de fonctionnement des stations d'épuration, au pas de temps journalier. Ces données sont stockées sur une base de données hébergée sur le cloud. Elle est spécialement conçue pour gérer les données temporelles. De plus, elle peut également s'installer en local sur une machine.

Les données sont téléchargées à chaque ouverture de l'application.

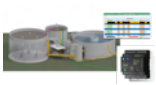
❖ *L'API*



Cette interface de programmation applicative également stockée en ligne permet la réalisation de différents services en continu.

Par exemple, une analyse sur le fonctionnement de chaque station est effectuée quotidiennement à partir des schémas de process définis et des prédictions des modèles d'IA. En cas de besoin (ex : prévision de dégradation du rejet), l'API envoi une notification sur les applications des utilisateurs liés à la station d'épuration et ajoute une tache qui doit être effectuée pour améliorer le fonctionnement. L'API permet également d'émettre des alarmes en cas de dépassement de valeurs seuils.

❖ *La STEP*



Afin de faciliter la tâche de l'exploitant, le recueil des données de fonctionnement des stations peut être automatisé. Un dispositif de connexion à l'automate de la STEP peut être mis en place et lire les données directement issues des capteurs

existants. Les données peuvent également provenir d'un fichier de relève rempli par l'exploitant (ex : Excel). L'API lit ce fichier à intervalle régulier et remonte les nouvelles valeurs sur la base de données distante.

b. *Connexion et navigation*

L'accès à l'application est sécurisé. L'utilisateur doit s'identifier à l'aide d'un nom d'utilisateur et d'un mot de passe fourni.

Afin de faciliter le processus, l'utilisateur peut utiliser les dispositifs d'identification sécurisés et intégrés des appareils IOS comme la reconnaissance faciale ou les empreintes digitales.



Page de connexion



Page de connexion

L'application est organisée en vues principales, permettant à l'exploitant d'accéder rapidement aux informations qu'il cherche.

❖ *Vue traitement*

Elle permet la visualisation du traitement avec le synoptique de traitement (ou une photo) et les valeurs enregistrées.

❖ *Vue Résultats*

Elle permet la visualisation des principaux indicateurs de suivi du traitement et de leurs niveaux.

❖ *Vue Mesures*

Les saisies des valeurs mesurées par l'exploitant se font sur cette vue. L'utilisateur peut choisir de rechercher par ouvrage la mesure ou suivre des feuilles de saisies paramétrables.

❖ *Vue Guide*

Cette vue regroupe les diverses analyses et visualisations sur le fonctionnement de la station. On y retrouve les courbes de traitement, les prédictions sur les paramètres de sortie, les conseils de réglages et les analyses du process épuratoire.

❖ *Vue Config*

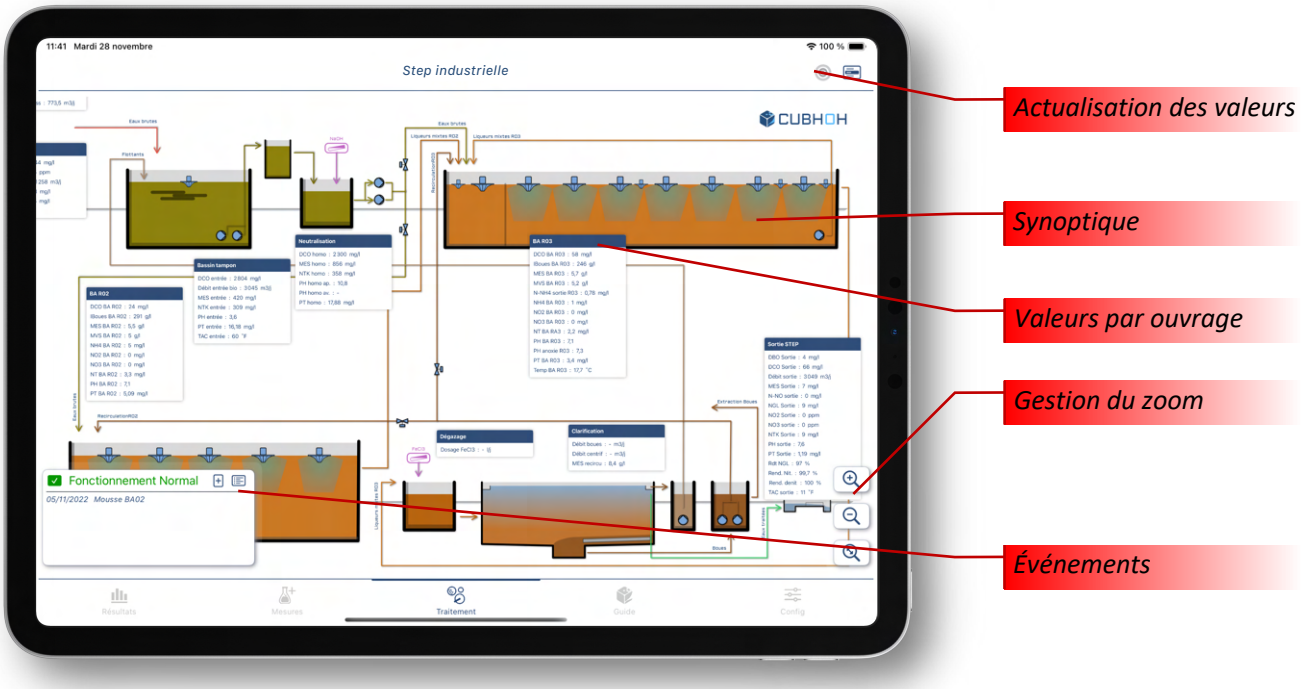
Cette vue permet principalement de sélectionner la station d'épuration que l'on veut visualiser. Les exports de données (Excel, Csv) sont également accessibles depuis ici. Pour les utilisateurs ayant les droits « administrateur », le paramétrage de tous les éléments composant la station peut être modifié à partir de cette vue.

Dans chaque vue, un menu permet d'avoir accès à des informations complémentaires, des actions ou des réglages.

Les détails des différentes vues sont décrits ci-après.

c. *Vue Traitement*

La vue présente, sous forme de synoptique de traitement ou de photo, les différents ouvrages de la station avec les dernières valeurs enregistrées.



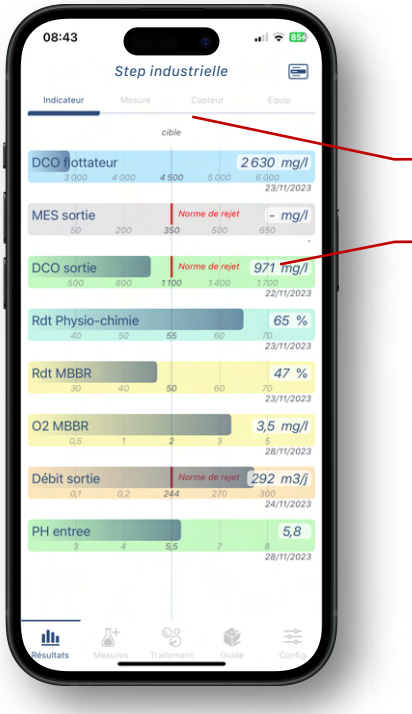
Cette page assure la visualisation synthétique des étapes de traitement, incluant les paramètres principaux suivis, actualisés au jour le jour. Un « clic » sur chaque ouvrage permet d'ouvrir la liste des paramètres.

Il existe 6 types de données dans l'application. Chaque type permet des fonctionnalités différentes.

- Les données « mesure » sont des données saisies manuellement par un exploitant (ex : mesure de DCO, valeur instantanée d'O2, MES …).
- Les données « capteur » sont celles récupérées automatiquement depuis un capteur connecté (via l'automate ou par de l'IOT).
- Les données « compteur » permettent de saisir des valeurs de totalisateur.
- Les données « réglage » permettent de saisir des valeurs de réglages de la station (ex : dosage de coagulant, mode de régulation de l'aération). Ce type permet de saisir des valeurs d'heure en plus de la date. De plus, la valeur de réglage est conservée jusqu'à la saisie d'une nouvelle valeur.
- Les données « calculé » sont générées automatiquement à partir d'autres données (ex : le rendement en DCO du flottateur en fonction de l'entrée et de la sortie).
- Les données « utilité » permettent de saisir un état d'un consommable ou de stockage de sous-produit. (ex : niveau de polymère, volume de boues primaires stockées).

d. *Vue Résultats*

La vue résultats permet d'afficher les indicateurs du fonctionnement de la station et également les meures des capteurs et des équipements.

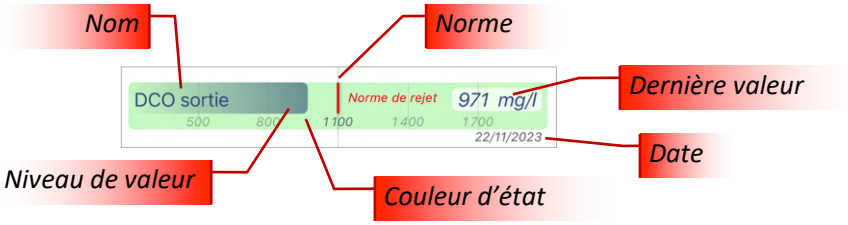


Cette vue affiche l'ensemble des indicateurs paramétrés. Un indicateur permet de visualiser facilement une donnée représentative de l'état de fonctionnement de la station.

Vues secondaires

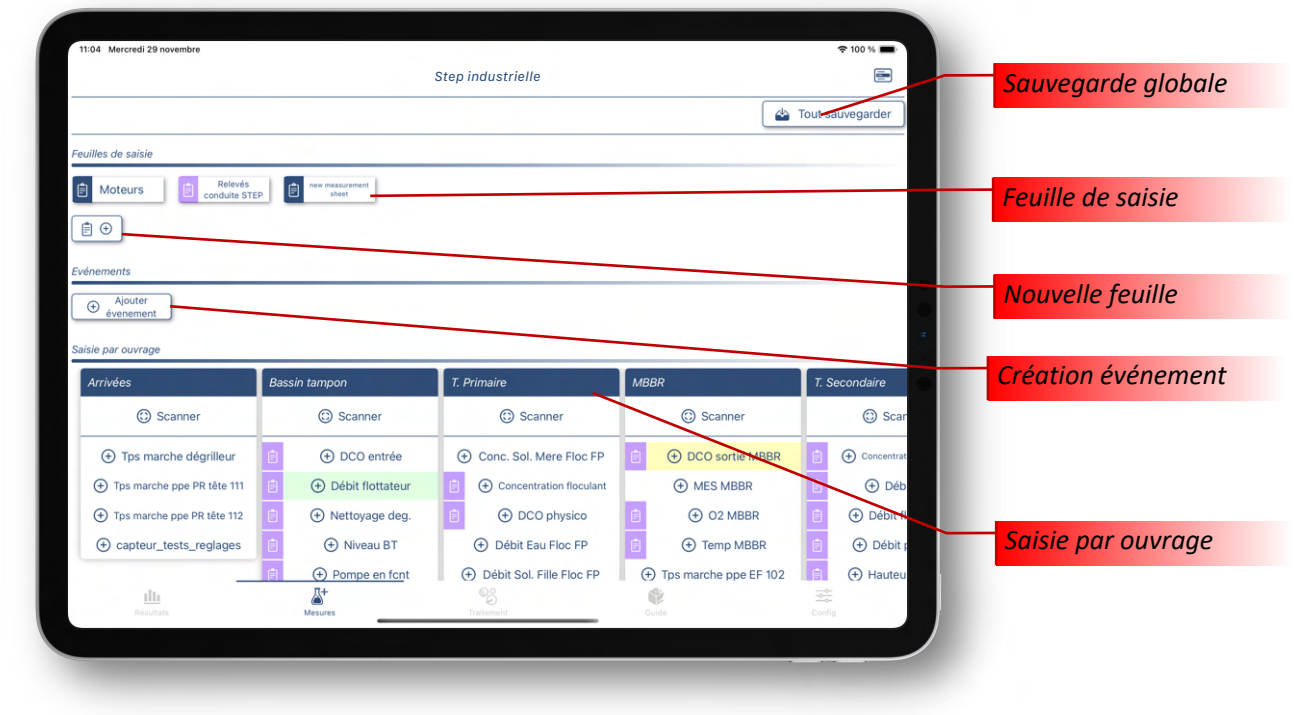
Indicateurs

Un code couleur est associé à la valeur de la donnée allant du bleu indiquant une conformité à l'orange alertant sur un dépassement.

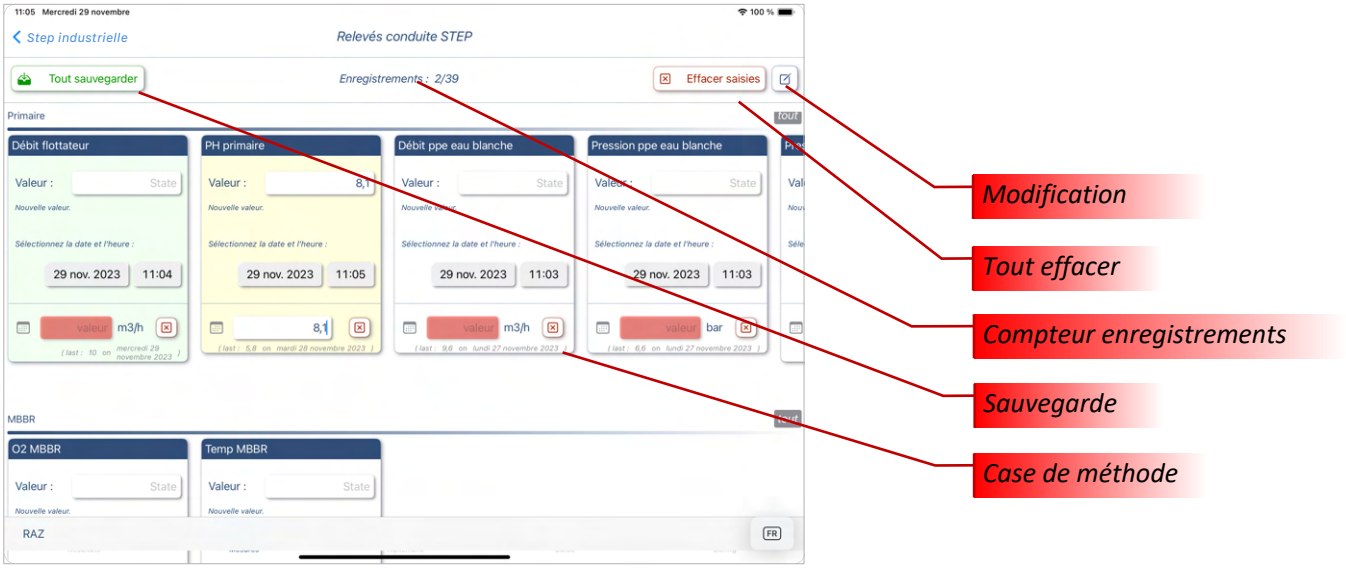


e. *Vue Mesures*

La vue présente l'ensemble des données pouvant être ajoutées.



Parmi ces données, la feuille de saisie a pour objectif de remplacer la feuille papier de relève que doit remplir l'exploitant lors de ses passages sur une station d'épuration. Configurable, elle permet à l'opérateur de disposer d'une feuille de route sur ses analyses et relevés à effectuer.



Chaque case de méthode varie en fonction du type de donnée.

f. *Vue Guide*

La vue guide donne accès aux différents conseils, analyses et prédictions que peut fournir l'application.

La vue est décomposée en 4 sous vues accessibles par le sélecteur en haut de l'écran.



f-1. A faire



Cette vue liste les éléments « A faire » relatifs à la station d'épuration. Ces éléments peuvent être générés automatiquement par notre API ou manuellement par l'utilisateur.

Ils peuvent concerner par exemple :

- Une alarme prédéfinie sur un capteur ou une mesure
- Une opération d'entretien à effectuer
- Une information sur des nouvelles prévisions sur les paramètres de sortie
- Une anomalie sur des données saisies
- Le résultat d'une analyse du process épuratoire automatique qui demande une action.
- …

f-2. Analyses process

Cette vue regroupe les différentes analyses qui sont effectuées sur le process épuratoire de la station d'épuration. Elles peuvent être sous forme de graphique ou de schémas automatisés.



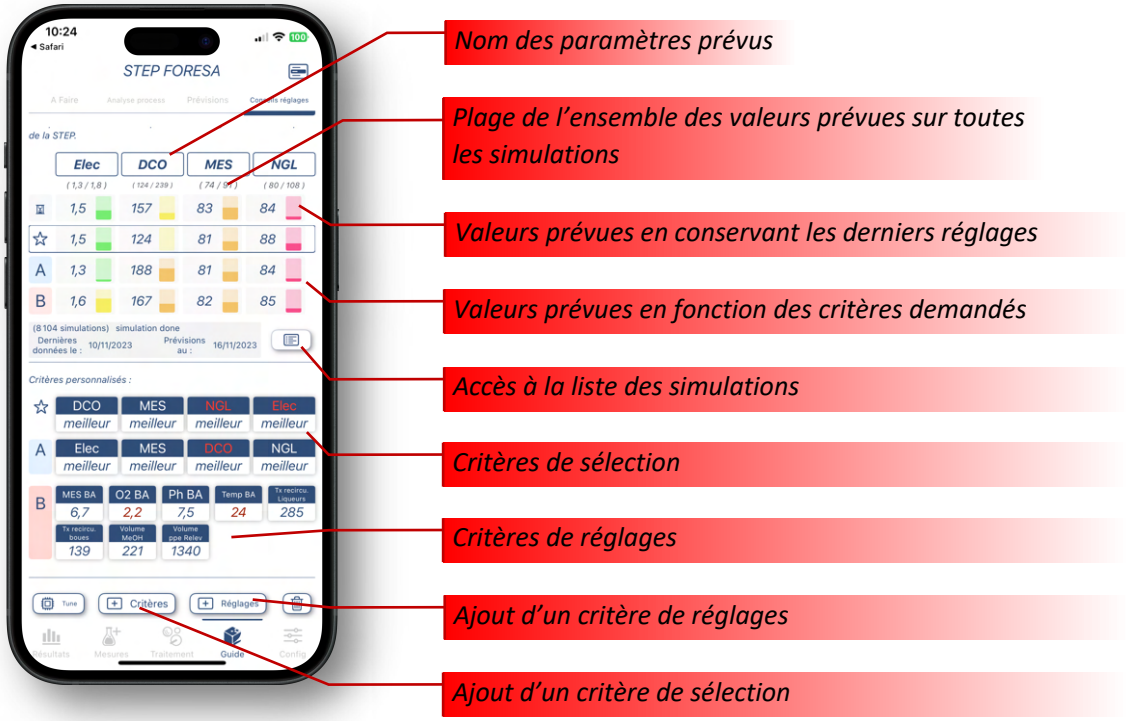
f-3. Prévisions

Cette vue regroupe les différentes prévisions sur les valeurs des paramètres de sortie en fonction des modèles d'IA. Chaque élement represente un paramétre.



f-4. Conseils réglages

Cette vue permet de donner des conseils de réglages en fonction du choix de l'utilisateur. L'utilisateur peut également modifier les régalges directement et visualiser les résultats prévus.



4. Modélisation numérique des unités de traitements

Une des fonctionnalités de l'outil est la prédiction des valeurs des paramètres en sortie de traitement et également la fourniture de conseils sur les réglages à adopter pour les améliorer. Pour permettre cela, nous avons voulu construire des modélisations numériques des STEP. Leur but est de pouvoir simuler le fonctionnement biologique d'une station et donc d'être capable de fournir les caractéristiques de l'effluents de sortie en fonction de celui d'entrée et des conditions de fonctionnement.

Pour réaliser cela, nous avions besoin d'outils permettant l'acquisition de donnée de fonctionnement, nécessaires à la modélisation, et utilisables ensuite pour tester et valider ces modèles.

a. *Pilotes de traitement*

Nous avons donc choisi de construire des unités pilotes de traitement des eaux. Le suivi quotidien de ces unités de traitement, à échelle réduite, nous a permis d'acquérir des données de fonctionnement. Leur exploitation nous a également servis de base pour concevoir notre application. En effet, nous avons assuré l'exploitation de ces unités et nous avons donc pu appréhender les besoins et les fonctionnalités nécessaires pour un outil d'aide à l'exploitation.

De plus, ces pilotes nous ont permis de développer les fonctionnalités de communication de capteurs de traitement (sondes, moteurs…) avec notre système applicatif (fonctionnalité IoT).

Nous avons construit 3 pilotes de traitement.

- *Un pilote « Agro »*

Ce pilote a reproduit une station d'épuration d'effluents vinicole classique, de type méthanisation + boues activées pour le traitement de la pollution organique. L'alimentation a été composée d'effluents synthétisés in situ pour cadrer les variations de qualité.

- *Deux pilotes « FORESA »*

Ces deux pilotes étaient les jumeaux miniatures de la station d'épuration du site de FORESA France. Pour valider les données fournies par les pilotes pour nos modélisations, nous avons dû nous assurer de leur fiabilité. Un pilote a donc servi d'étalon pour caler la reproductibilité de la station réelle et un autre nous a permis de simuler et d'entraîner des variations de fonctionnement (alimentation, paramétrage, incidents…)

L'alimentation était réalisée par l'amenée des effluents de FORESA 1 à 2 fois par semaine par nos soins (réaliste par la capacité du bassin tampon de FORESA qui représente entre 10 et 15 jours de temps de séjour).

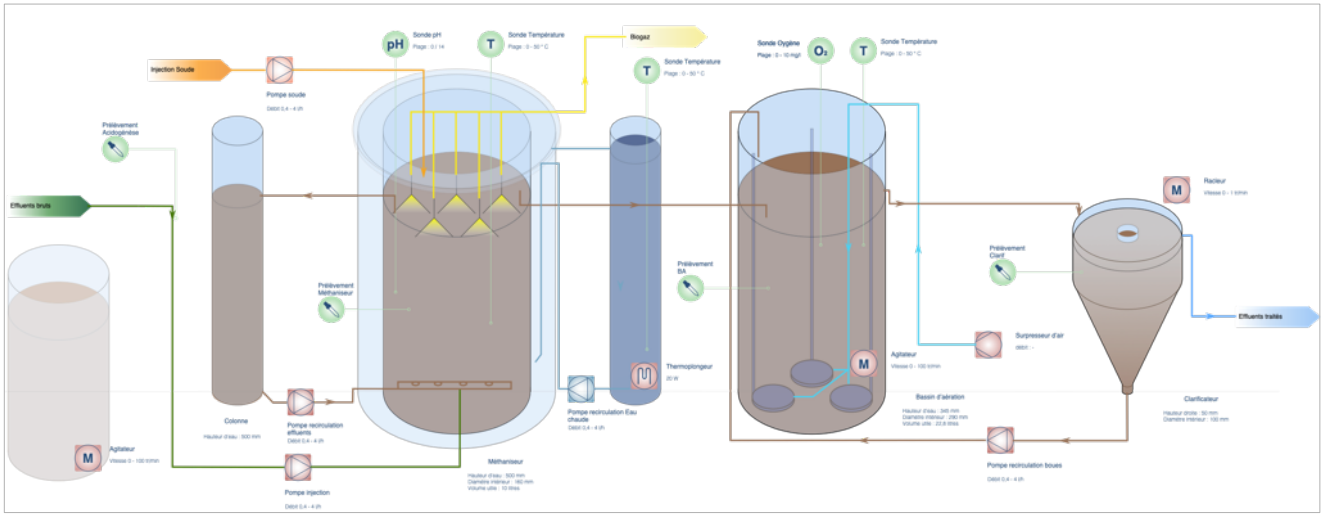
a-1. Construction

❖ *Design du traitement*

Le pilote « agro » était composé des ouvrages suivants :

- Un bac d'homogénéisation des effluents,
- Un méthaniseur et sa colonne d'équilibrage/ recirculation, de même qu'un système de réchauffage par circulation d'eau chaude dans une double enveloppe (circuit thermostaté secondaire),
- Un bassin d'aération
- Un clarificateur

Le schéma suivant détaille la conception. Il indique les ouvrages et les équipements et également les points de mesure pour le monitoring du fonctionnement.



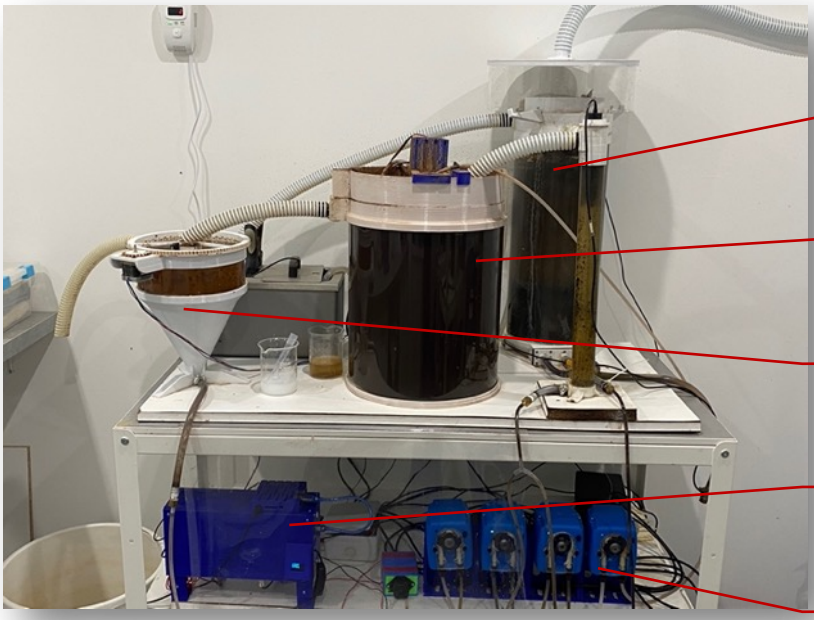


Photo du pilote « agro »

Les pilotes « FORESA » étaient composés des ouvrages suivants :

- Un bac d'homogénéisation des effluents,
- Un bassin d'anoxie,
- Un bassin d'aération
- Un clarificateur

Le schéma suivant détaille la conception. Il indique les ouvrages et les équipements et également les points de mesure pour le monitoring du fonctionnement.

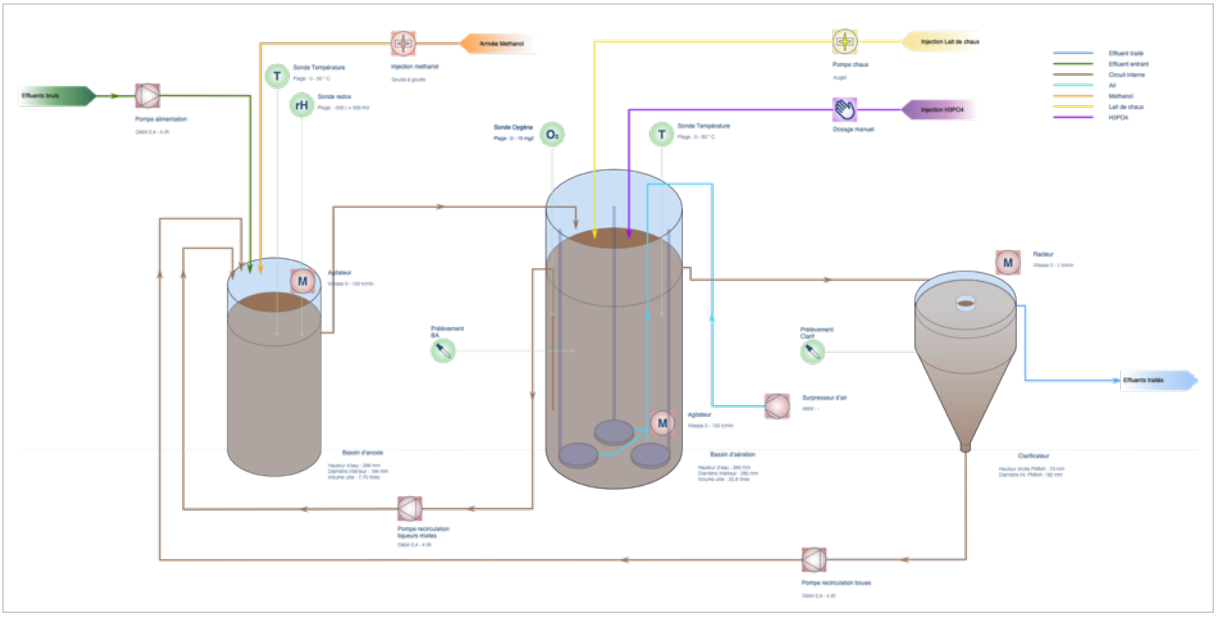




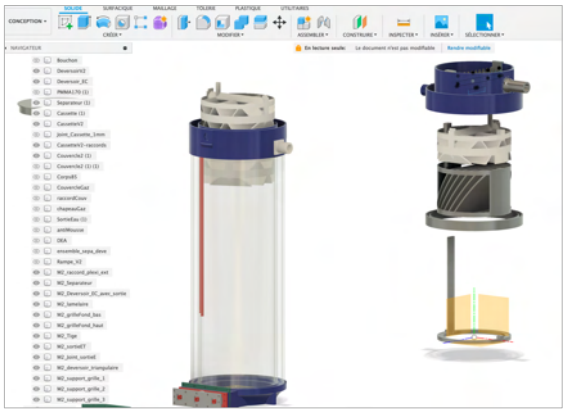
Photo des pilotes « FORESA »

❖ *Fabrication*

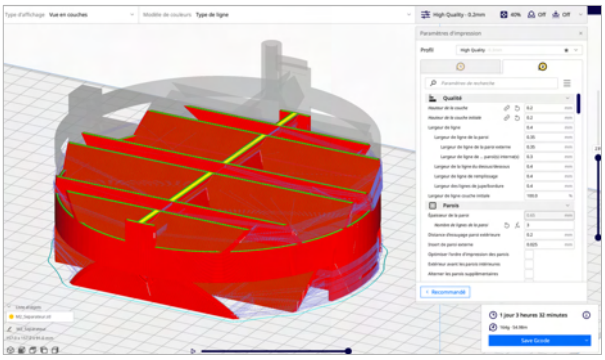
Nous souhaitions que les pilotes soient facilement adaptables afin de résoudre les problèmes éventuels d'échelles entre ces unités de laboratoire et les stations à tailles réelles. Nous avions également besoin d'avoir la possibilité de modifier le design épuratoire pour tester différents dimensionnements.

Pour cela nous avons maitrisé l'ensemble du processus de fabrication en construisant nous même les pilote grâce à de l'impression 3D additive.

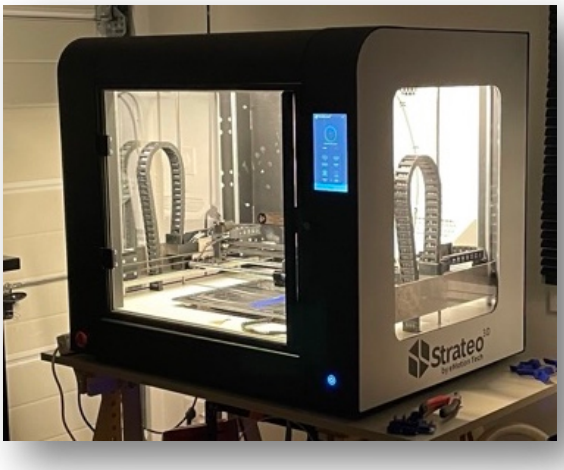
Les éléments ont été modélisés en DAO et ensuite imprimés en 3D.



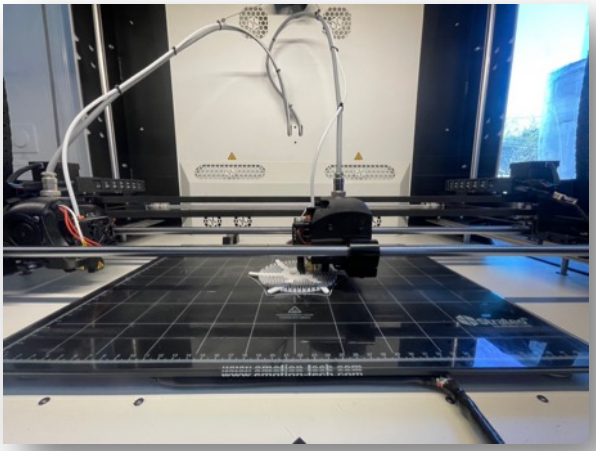
Conception sous DAO



Préparation pour l'impression 3D



Imprimante 3D



Impression



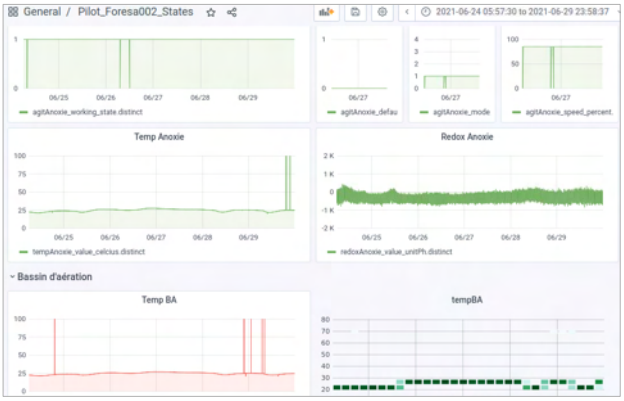
Détail des pièces

❖ *Contrôle commande*

L'automatisme des pilotes était assuré par un système informatique que nous avons conçu. Il assurait le pilotage des moteurs et intégrait la sauvegarde automatique de toutes les valeurs de fonctionnement sur des bases de données locale et distante.



Unité de contrôle

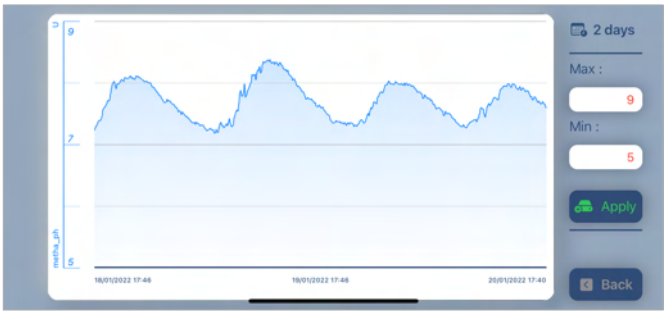


Archivage des données en local

Nous avons conçu une application supplémentaire pour piloter le fonctionnement. A distance, elle permettait de visualiser l'état des moteurs, de changer leur mode de fonctionnement, de réguler l'injection de soude pour maintenir le pH et de visualiser les valeurs des capteurs. Un système de notifications permettait d'être alerté en temps réel d'un dysfonctionnement sur les unités.

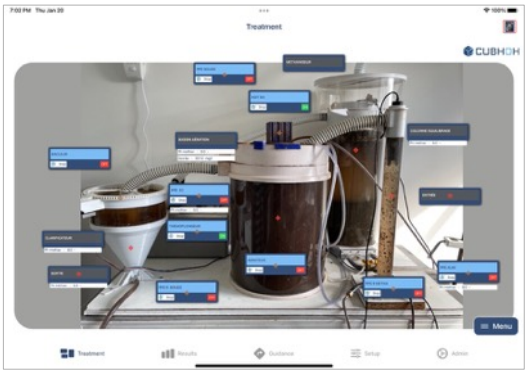


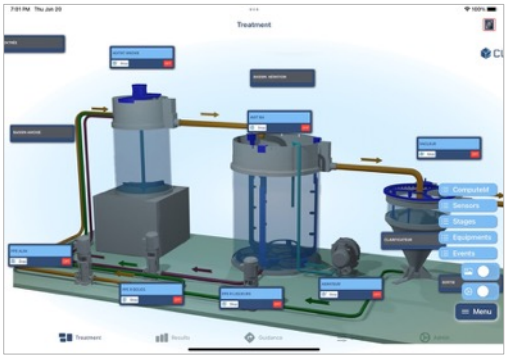




Vues de l'application ASTRHOH

Enfin, les mesures et analyses étaient saisies et suivi sur notre outil d'assistance en cours de développement.





a-2. Suivi des pilotes

Les 3 pilotes ont été exploités dans les locaux de SPEC Environnement ce qui a permis un suivi journalier des analyses et des paramètres de fonctionnement.

Des bilans hebdomadaires étaient réalisés par la confection et l'analyse d'échantillons moyens par un laboratoire accrédité COFRAC, pour les paramètres suivants :

- Entrée station : pH, DCO, DBO_5, MES, NTK, NH_4^+, NO_3^-, NO_2^-, NGL, Norga, Ptotal et TAC
- Sortie station : pH, DCO, DBO_5, MES, NTK, NH_4^+, NO_3^-, NO_2^-, NGL, Norga, Ptotal et TAC

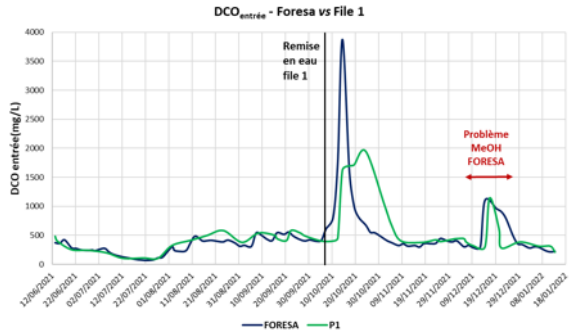
En complément un suivi quotidien était réalisé sur les paramètres complémentaires suivants :

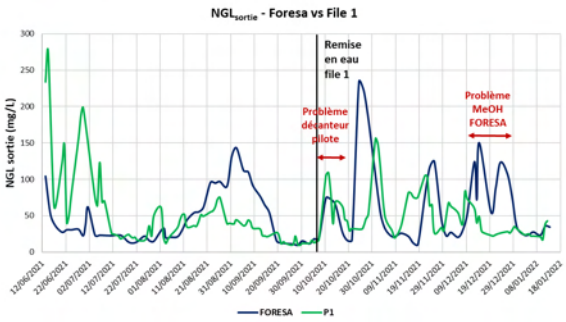
- Sortie méthaniseur pilote « Agro » : pH, TAC, AGV, DCO, NGL, PT
- Sortie BA pilotes « FORESA » : MES, MS, MVS, Indice de boues, O_2 , température, pH

Les essais se sont déroulés sur 1 année (juin 2021 à mai 2022).

a-3. Bilan

Le suivi des pilotes et les différents essais ont montré que les pilotes assurent une bonne reproductibilité des unités de traitement des eaux industrielles, comme le montrent, à titre d'exemple, les courbes suivantes (comparaison entre les résultats d'autocontrôle de la station réelle (FORESA) et de l'unité pilote (File 1) :





Cette reproductibilité validée, nous avons pu simuler divers scénarii sur le pilote File 2 en conservant le pilote File 1 comme étalon (fonctionnant à l'identique de la STEP).

Les unités pilote permettent ainsi de capitaliser de la data pertinente permettant d'alimenter les algorithmes d'IA.

b. *L'Intelligence Artificielle*

Le travail sur la modélisation des données par de l'IA s'est décomposé en deux principales phases. La première a été de faire la démonstration de la faisabilité et ensuite durant la deuxième de concevoir des modélisations plus fines et pouvant être utilisées dans notre outils.

b-1. Preuve du concept (PoC)

Pour cette étape, nous avons travaillé avec la société JOLIBRAIN, basée à Toulouse. Elle est composée d'experts en l'IA. Elle conçoit des services à haute valeur ajoutée qui intègrent les dernières avancées en matière d'intelligence artificielle pour les industriels.

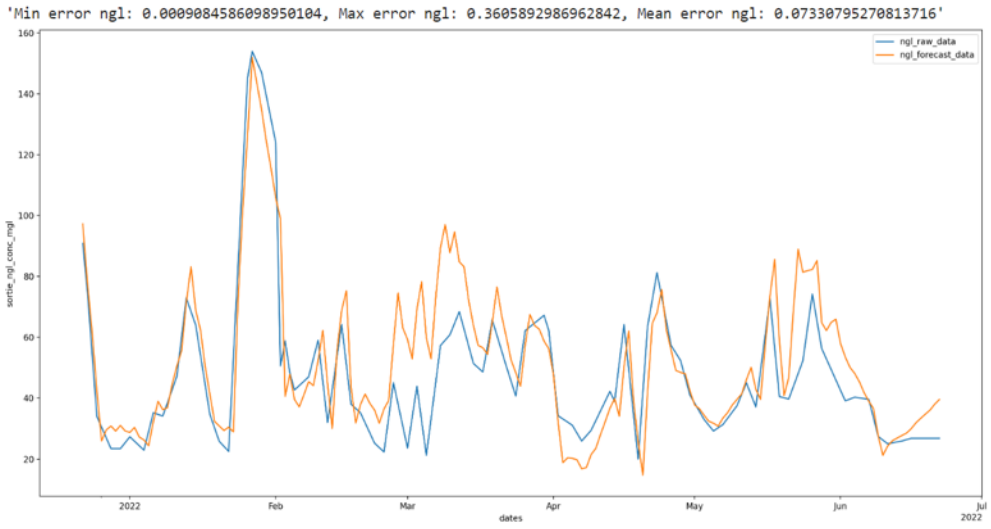
A partir de son expertise, JOLIBRAIN a choisi d'utiliser un modèle de réseau de neurones adapté aux prédictions de séries temporelles.

Le modèle comporte plusieurs éléments à configurer :

- Les données d'entrées (données de fonctionnement de la station que l'on va prendre en compte pour les prédictions)
- Les données de sorties (le ou les paramètre de sortie du traitement que l'on souhaite prédire)
- La distance en jour de la valeur à prédire,
- Les hyperparamètres du modèle et de l'entrainement

Différents entrainements ont été effectués en prédisant la sortie de J+1 à J+14.

A titre d'exemple, les résultats obtenus à J+7 sont présentés ci-dessous (comparaison entre les valeurs réelles en NGL en sortie station et les valeurs issues de l'IA):



La démonstration du concept est donc validée : l'IA permet une modélisation numérique capable de prédire, à un horizon de 7 jours environ, les données de sortie d'une station d'épuration en fonction de l'historique des données.

b-2. Validation avec les pilotes

A partir des données historiques enregistrées depuis 2 ans sur l'usine de FORESA et de celles enregistrées sur les pilotes durant les 12 mois de suivi, nous avons pu affiner notre travail sur la modélisation.

L'objectif des nombreux essais d'entrainement a été d'apprécier l'apport des données issues des pilotes. Ceci afin de modéliser un système de traitement avant sa construction ou d'affiner sa précision avec des cas d'usages difficilement réalisables sur une unité en production.

La modélisation n'est pas parfaite, mais le modèle a « appris » grâce aux données issues des essais des pilotes.

b-3. Choix des outils de modélisation

La démonstration de la faisabilité de modéliser numériquement le traitement d'une station en y incorporant des données issues de pilotes faite, nous avons voulu faire valider cette démarche par un spécialiste externe.

Nous souhaitions une analyse critique de notre démarche pour la valider ou non et connaitre ensuite les autres outils disponibles avec leurs performances.

❖ *Intervenant externe*



Nous avons choisi Vincent BERIONNI. C'est un Data Scientist spécialisé dans le domaine du traitement de l'eau. Il a été employé chez SAUR à Maurepas. Il a pu notamment travailler sur l'élaboration des modèles prédictifs suivants :

- Défaillance sur le réseau d'eau potable,
- Débits horaires d'effluents entrant sur une station d'épuration en fonction des prévisions de précipitations,
- Niveau de la ressource en eau dans les forages

Nous lui avons présenté notre méthodologie et nous lui avons transmis le travail déjà réalisé. A partir de cela, il nous a transmis son avis. Il a également pu intervenir lors d'une réunion du COPIL.

Son avis est présenté ci-dessous.

La solution mise en œuvre utilise la plateforme DeepDetect, développée par la société Jolibrain.

Cette plateforme s'appuie notamment sur le framework de machine learning Pytorch, qui est l'un des plus aboutis et les plus utilisés par la communauté du deep learning, avec les autres grands acteurs de la discipline que sont Keras et Tensorflow.

Plus précisément, le choix s'est porté sur modèle qui se base sur des réseaux neuronaux pour développer une orientation spécifique à la prédiction de séries temporelles. Cet algorithme est l'une des solutions IA les plus en pointe actuellement existante pour la construction de modèles de séries temporelles, c'est-à-dire adapté à la problématique de prédiction de l'évolution des variables en sortie de station d'épuration. Ce modèle étant toutefois gourmand en temps et ressources machine, il pourra notamment faire l'objet d'optimisations dans le choix des hyperparamètres qui guident la calibration du modèle, ou être confronté à d'autres modèles plus légers et véloces. Ceci de manière à déterminer si le gain de temps passe ou non par une réduction rédhibitoire de la performance du modèle.

La méthode d'évaluation de la performance du modèle est standard, avec train, validation et test. Elle pourrait être renforcée à l'avenir par une méthode dite de « forward cross validation » qui permet de mieux stabiliser le modèle.

Par ailleurs, la visualisation de courbes ou d'une zone matérialisant un intervalle d'incertitude présenterait l'avantage de pouvoir définir un degré de confiance dans le résultat, et fournir une meilleure « explicabilité » des écarts qui peuvent parfois subvenir entre prévu et réel.

Le pas de temps de ré-entraînement prévu, une fois par mois, semble cohérent et correspond à ce qui se pratiquait pour des problématiques similaires d'après mon expérience.

D'autres travaux d'exploration pourraient amener à perfectionner le modèle, le rendre plus rapide, ou encore aider à une interprétation plus humaine compréhensible des résultats, en évitant ainsi l'effet « boîte noire » reproché à l'IA et notamment aux modèles de type réseaux de neurones.

L'approche qui consiste à entraîner un modèle sur le pilote pour l'appliquer ensuite en prédiction à la STEP est à ma connaissance novatrice en ce qui concerne les stations d'épuration. Je pense qu'il y a un grand intérêt à approfondir (dans la mesure ou ça permet vraiment de se démarquer d'autres solutions existantes), notamment pour améliorer la normalisation / calibration qui permet de passer de l'un à l'autre et vice-versa.

L'étape de recommandation se base actuellement sur une approche classique (mais sûre) de type « grid search », c'est-à-dire exploration de tout l'espace des paramètres. Elle pourra faire l'objet dans un second temps d'une optimisation à l'aide de méthodes plus poussées, qui permettront là aussi de réduire le temps machine nécessaire.

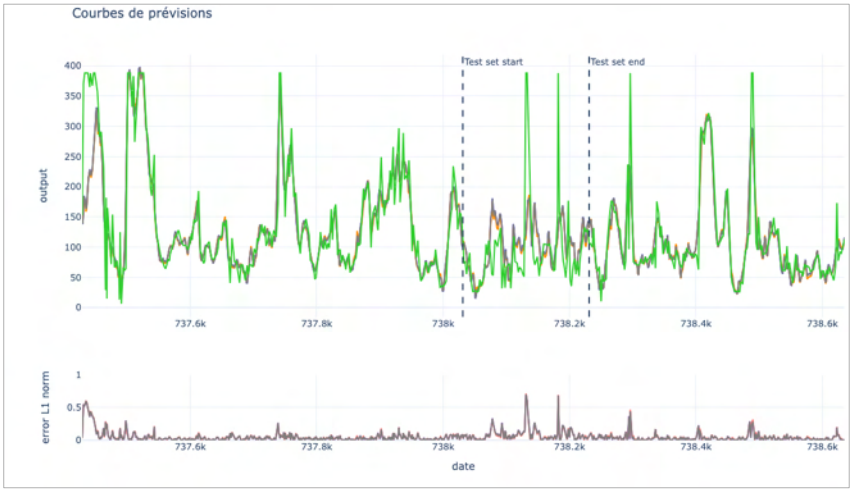
Cette analyse conforte le travail que nous avons accompli, et le choix du modèle utilisé, et nous offre quelques pistes d'amélioration et de travail.

b-4. Mise en production

A partir du choix retenu précédemment, nous avons donc utilisé la librairie Pytorch pour concevoir les modèles des stations d'épuration utilisés dans notre application.

Pour nous permettre de convertir les modèles et d'utiliser les outils complémentaires, nous avons donc utilisé des scripts directement en python pour entrainer nos modèles en Pytorch et sortir de la plateforme de DEEPDETECT.

La figure ci-contre illustre, une fois de plus, la pertinence du modèle en phase de production.



5. Le déploiement

a. *Préambule*

Le principal objectif de cette phase a été de vérifier l'adaptabilité de l'outil, et la célérité de mise en œuvre sur des sites différents avec des stations d'épuration non-similaires. En effet, l'architecture de l'outil s'est voulue modulaire, aussi bien dans la définition du process épuratoire que dans la gestion des bases de données associées. De cette façon, les outils annexes (mesures, reporting, modélisation, prévisions, conseils…) sont communs et sont d'ores et déjà fonctionnels.

Dans les 4 cas étudiés, en plus du site initial de FORESA, cette adaptabilité a été vérifiée. La création de chaque « feuille de bord » pour chaque station ne prend que quelques heures d'assemblage de blocs prédéfinis (ouvrages, équipements et paramètres associés). A chaque nouvel ouvrage ou équipement rencontré, la création de celui-ci et de ses attributs (paramètres, mesures…) ne prend que quelques heures de plus et s'ajoute à la base de donnée générique permettant d'être réutilisé à l'avenir sur d'autres sites.

Le deuxième objectif de ce déploiement a été de parfaire certaines fonctionnalités ou de vérifier le bon fonctionnement de celle-ci sur des sites différents. Cependant, chaque site n'a pas bénéficié de toutes les fonctionnalités, car cela n'était pas pertinent et/ou trop chronophage dans le cadre de cette partie de projet. Par exemple, la modélisation par IA n'a pas été regardée sur les sites ne disposant pas de suffisamment de données historiques. Ou encore, la connexion de l'automate à l'application n'a été regardée que sur les sites qui l'ont permis.

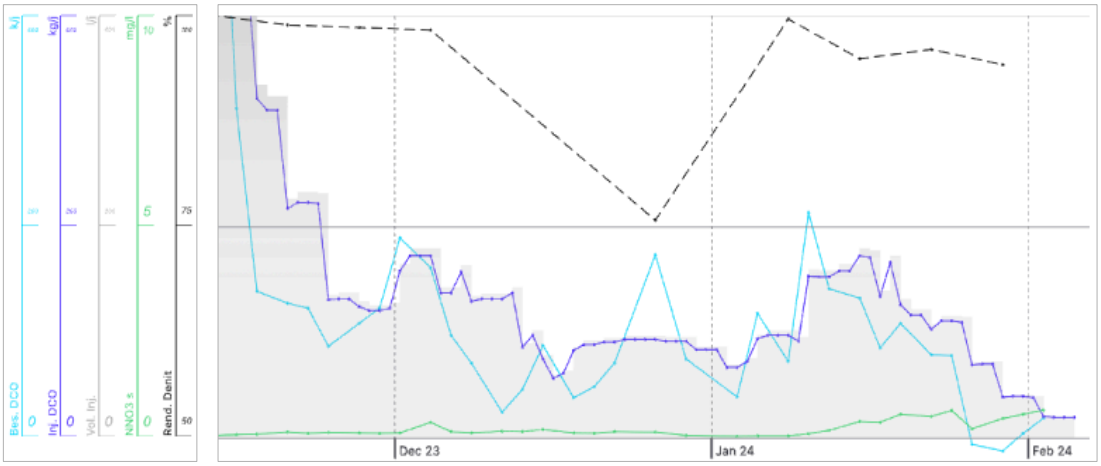
Le troisième objectif a été de pouvoir organiser, à l'échelle d'un administrateur multisites, les différents interfaces de plusieurs stations d'épuration sur un seul et même outil. Cette fonctionnalité est particulièrement pertinente pour une société de conseils ou de service gérant cette configuration.

b. *FORESA*

L'entreprise est un industriel de la Chimie à Ambarès-et-Lagrave (33). Elle produit des colles formol-urée.

Outre la capitalisation des données et le partage des données de suivi de la station, l'utilisation de l'application a en particulier permis :

- De réduire les consommations en réactifs, et en particulier, le méthanol qui est ajouté en fonction des charges en azote et en DCO entrantes sur la station. L'application permet d'assurer un suivi journalier des besoins en méthanol calculés sur la base de la DCO et de l'azote entrants. Le graphe ci-dessous, extrait de l'application, présente :
    - En bleu clair, le besoin calculé en méthanol (calculé au jour le jour),
    - En violet, l'apport réel apporté en DCO grâce au méthanol,
    - En gris la quantité de méthanol apportée.



La période considérée montre que l'ajustement n'est pas fait au jour le jour pour le dosage en méthanol (par manque de suivi de l'opérateur). Le cas échéant, l'ajustement quotidien du méthanol permettrait un gain de 30 litres /jour, soit une économie potentielle de 12 €/jour, soit environ 1000 € sur la période considérée de 90 jours.

- De réduire les non-conformités des rejets en sortie de station, en augmentant la maîtrise du process. Sur ce point, les conseils prodigués par l'application, issus des prédications générées par l'IA, assurent un réglage plus fin des paramètres de traitement (débits, pH, nutriments…).
- On peut en particulier noter l'incidence positive des simulations faites pour l'élimination de l'effluent « POI » limité à 5% du débit entrant. Au-delà, le rejet en sortie de station aurait été non-conforme. On peut noter de plus, le gain financier pour ne pas externaliser l'élimination de ces flux supplémentaires. Le volume à traiter étant d'environ 5000 m^3, le coût d'externalisation peut-être compris entre 100 et 500 k€.

C. *Industriel n°2*

L'industriel n°2 est spécialisé dans le secteur du bois. La station d'épuration est composée :

- D'un bassin tampon
- D'un premier étage physico-chimique de floculation/flottation,
- D'un étage intermédiaire de type biologique forte charge,
- D'un étage secondaire de traitement biologique de type faible charge,
- D'un clarificateur,
- D'une filtration tertiaire avant rejet au milieu naturel.

L'intérêt du partage des données de suivi de la station est encore plus marqué sur ce site que sur FORESA, car l'exploitant principal est le personnel interne de l'usine (pour rappel, prestation sous-traitée chez FORESA). Par conséquents, les intérêts d'utilisation de l'application sont les suivants :

- Montée en compétence des opérateurs non-initiés au traitement des eaux,
- Reporting en direct et conseils à distance (« hotline » SPEC Environnement),

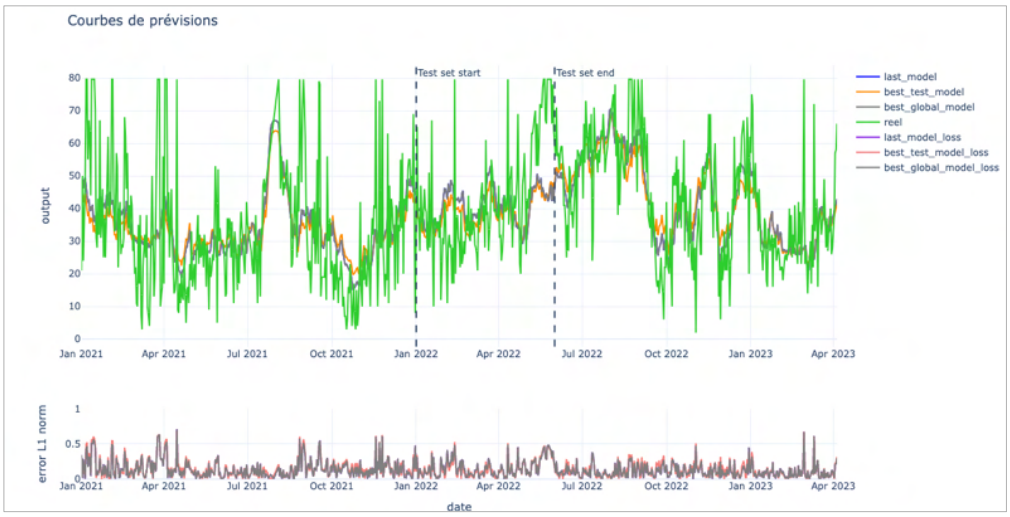
- Optimisation des injections des réactifs par des calculs incrémentés dans l'application (ajouts d'urée et d'acide phosphorique en fonction de la DCO brute et des rendements du premier étage physico-chimique). Sur la station, les couts de réactifs relatifs à cette optimisation ne représentent que quelques centaines d'euros par an. Par contre, la non-maîtrise de ces dosages peut entraîner des non-conformités sur les rejets, avec des conséquences financières potentiellement très impactantes pour l'entreprise (arrêt de la production pour interdire les rejets avec une diminution du chiffre d'affaire de 80 à 100 k€/ jour, ou élimination en centre extérieur (estimation à environ 10 k€/jour supplémentaire)).
- Gestion des stocks de réactifs avec alarme,
- Optimisation de l'aération et donc gains sur l'énergie dépensée. A titre d'exemple, l'apparition de nitrates en sortie de station a été suivi grâce à l'application ce qui a permis d'aménager des plages d'anoxie pour permettre la dénitrification, engendrant des phases de non-aération de l'ordre de 20% du temps, soit une économie du même ordre sur la dépense énergétique de ce poste (économie estimée à environ 2 k€/an).
- Réduction des charges de sous-traitance externe par internalisation de l'exploitation. Après une année de mise en œuvre de l'application, le coût de sous-traitance externe a été réduit de moitié, soit un gain d'environ 25 k€/an).

d. *Industriel n°3*

L'industriel n°3 est spécialisé dans l'agro-chimie (extraction de matières actives sur des produits d'origine animale).

La station d'épuration est de type biologique dimensionnée pour le traitement du carbone et de l'azote. La filière est complexe avec 2 étages de boues activées (1 chenal d'aération et 1 bassin en alternance de phases) en parallèle mais interconnectés.

Les données de suivi sur cette station sont nombreuses et ont permis de construire un modèle d'IA spécifique en utilisant l'architecture développée. Le graphe ci-dessous présente les résultats de la simulation comparée aux valeurs réelles, et montre ainsi la pertinence évidente de ce modèle :



e. *Industriel n°4*

L'industriel n°4 est un site agroalimentaire dans le domaine du lait.

La station d'épuration est composée :

- D'un bassin tampon
- D'un premier étage physico-chimique de coagulation/floculation/flottation,
- D'un étage secondaire de traitement biologique de type forte charge,
- D'un flottateur final avant rejet au réseau communal.

Contrairement aux autres sites, la mise en place de l'application a été concomitante aux travaux de revamping de l'installation. Par conséquent, le développement la partie prédictive n'a pas été réalisée, en raison du faible volume de data. L'intérêt de l'utilisateur (exploitant interne du site) a été fortement porté sur 2 aspects :

- Archivage, mise en page et corrélation des différents paramètres de suivi pour un reporting facilité, avec imports/exports possibles via Excel. Cela lui permet d'aller au-delà de ce que lui propose la supervision mise en œuvre sur la station.
- En cas de non-conformité, la collectivité peut interdire le rejet des effluents dans le réseau communal. La conséquence immédiate est donc l'arrêt de l'usine ce qui représente une perte de chiffre affaires d'environ 200 k€/ jour.
- Connexion des capteurs en direct avec reports d'alarme sur l'application. Cette mise en place permet d'assurer naturellement une augmentation de la maîtrise de l'outil épuratoire, et une réduction du temps de présence de l'opérateur et des coûts associés.

f. *Industriel n°5*

L'industriel n°5 est également un site agroalimentaire dans le domaine du lait.

La station d'épuration est composée :

- D'un bassin tampon
- D'un étage biologique faible charge de type SBR (bassin séquencé assurant l'aération et la décantation) avant rejet au milieu naturel.
-

La typologie de traitement étant plus simple que sur les autre sites étudiés, l'exploitation est très succincte, et le suivi est plutôt « léger ». Ainsi, l'intérêt de l'application est moins évident que pour les autres sites. Néanmoins, on peut noter :

- La capitalisation des données en cas de changement d'exploitant interne. Les feuilles manuscrites sont reportées sous formes de courbes et de données archivées permettant de garder une mémoire de la vie de la station,
- la prise de conscience d'élargir le panel des indicateurs de suivi à d'autres paramètres, qui trouvent dans l'application, une explication documentée (suivi de l'azote et du phosphore, suivi des consommations en réactifs…).

g. *Enseignements des déploiements réalisés*

g-1. Typologie de traitement

L'objectif initial de développement de l'outil était de pouvoir disposer d'une application facilement déployable quelle que soit la filière de traitement. Bien entendu, l'intérêt de la fonction prédictive à 6 ou 7 jours est particulièrement intéressante quand un (ou plusieurs) étage(s) est de nature biologique. En effet, pour des filières purement physico-chimiques, les temps de réaction/impact suite à des changement des nature des effluents, ou « d'incident » sur le process sont beaucoup plus courts.

Au vu des quelques cas évoqués ci-dessus, la présence d'un seul étage ou de plusieurs étages de traitement n'est pas une limite au déploiement de l'outil. La gestion « modulaire » de l'application et des bases de données associées, de même que la souplesse de l'algorithme d'IA, laissent ouvertes toutes les configurations possibles.

g-2. Quantité et qualité des données

L'outil a pour finalité un diagnostic en continu de ce qui se passe sur la station d'épuration sur lequel il est mis en œuvre. Plus la quantité de paramètres suivis est importante, plus la pertinence des prévision (process ou IA) est avérée, ce qui est, par nature, le cas pour tout diagnostic.

De même, la qualité des données collectées est indiscutablement primordiale là encore, pour rendre pertinents les résultats du diagnostic continu.

L'expertise process apportée lors de la mise en place de WATHOH-WATHOH est donc nécessaire pour :

- Valider la qualité des données collectées (tri des valeurs, vérification des capteurs, analyses croisées sur différents paramètres, cohérence des données entre elles…),
- Permettre un tri par l'outil, et éventuellement une correction/estimation des données manquantes,
- Assurer que le nombre de données, s'il est insuffisant dans l'historique de la station, puisse être comblé par une collecte suffisante après le déploiement de l'outil. Ces données nécessaires sont de 2 types :
    - Données de fonctionnement des équipements (temps de fonctionnement, capteurs…)
    - Données de suivi (analyses, observations, etc…).
- S'assurer de l'intégration du bon « pas de temps » pour l'ensemble des paramètres et analyses. L'outil déployé est configuré pour fournir des conseils à la journée, et prévoir des évolutions sur plusieurs jours.

g-3. Durée nécessaire au déploiement de l'IA

Comme présenté précédemment, la pertinence de l'algorithme d'IA s'améliore avec l'accumulation des données. L'outil est de plus en plus précis dans le temps car il a la faculté « d'apprendre » au fur et à mesure.

Non seulement la quantité et la qualité des paramètres sont importantes (cf. paragraphe ci-dessus), mais la durée de la capitalisation de ces données est également déterminante. Il a été observé, dans les différents déploiements, que ce soit à l'échelle pilote, ou à l'échelle industrielle, qu'une durée d'un an a minima commence à être pertinente.

g-4. Wathoh-Wathoh = outil de terrain

Wathoh Wathoh se veut être un outil adapté à chaque cas particulier. Ainsi, les analyse process, qu'elles soient issues de l'expertise métier des ingénieurs de SPEC Environnement/CUBHOH ou issues de l'IA ne sont pas des copier/coller d'un site à un autre.

De même, en l'absence de données, comme par exemple, à la construction d'une nouvelle station d'épuration, l'outil ne permet pas d'assurer une modélisation sur la base d'un modèle établi. Le recours au stade pilote est alors nécessaire, a minima, pour construire les bases de cet algorithme et le retour d'expérience pour l'analyse process.

Rappelons ici que la fonction prédictive liée à l'IA n'est qu'une fonctionnalité de l'outil. La pertinence de WATHOH WATHOH est également importante pour les fonctionnalités suivantes :

- Capitalisation des données sur un seul et même outil (analyses, données de fonctionnement, observations, alarmes…)
- Conseils prodigués permettant un « apprentissage » des exploitants moins aguerris,
- Visualisation « en un clin d'œil » de l'état de fonctionnement et de conformité de la station,
- Visualisation des données sur des durées d'historiques non-limitées, et possibilité de construire ses propres « graphes » pour corréler différents paramètres. Observation des dérives possibles.
- Partage des informations par simple installation de l'application sur des appareils mobiles (téléphone, tablette),
- Outil évolutif « ouvert » permettant l'intégration d'autres données, qu'elles soient propres à la station d'épuration ou en amont/aval de celle-ci (données de production, suivi hydrique de la rivière, etc…),
- L'architecture de l'outil étant commune pour tous les sites, l'évolution des fonctionnalités, grâce aux mises à jour, est de fait répercutée sur tous les sites utilisant l'application,
- Possibilité pour les exploitants multisites de visualiser plusieurs stations avec un même outil,
- Limitation des erreurs de saisie, en particulier en utilisant la fonction « scanner »,
- Gestion des stocks.

6. Synthèse du projet

Le projet WATHOH-WATHOH a permis de concevoir et de mettre en œuvre un outil d'aide à l'exploitation des stations d'épuration industrielles.

L'outil a été conçu pour être adaptable et être facilement déployable sur de nouvelles stations, ce qui a pu être vérifié lors du déploiement. Les fonctionnalités développées ont été conçues et adaptées aux besoins de l'exploitation.

L'outil permet de capitaliser les données de fonctionnement et de visualiser le process de traitement. Il facilite la saisie des analyses et des mesures sur site mais peut également être connecté aux automatismes existants. Grâce aux analyses process, l'exploitant dispose de conseils pour résoudre les dysfonctionnements. Ces guides sont orientés par de l'expertise en traitement de l'eau et prennent appui sur les paramètres de fonctionnement réels. De plus, la modélisation du traitement par des algorithmes d'IA est possible. Son utilisation permet d'indiquer à l'exploitant les tendances de fonctionnement de son traitement dans le futur. Il peut orienter ses choix d'exploitation en fonction des prévisions du système. L'application lui prodigue des conseils avec des scénarios de réglages adaptés pour améliorer le traitement.

L'application a pu être utilisée par les exploitants. Elle a montré qu'elle pouvait être source d'économie de temps dans le suivi de l'exploitation, et une source d'économie également dans la gestion des réactifs et de l'énergie. Le conseils adaptés pour améliorer et pérenniser le traitement assurent enfin et surtout **une meilleure maîtrise du traitement et limite ainsi les non-conformités et l'impact sur le milieu naturel**.

La reproductibilité des unités de traitement par les pilotes a été démontrée. Ces derniers permettent d'acquérir des données complémentaires pour les modèles d'IA mais ils apportent également une base d'expérience puissante capable de conforter le niveau du traitement. Le facteur d'échelle mécanique n'impacte pas la reproductibilité du traitement. Il peut être réduit par la construction de pilote à façon facilement adaptables par l'impression 3D.

Nous continuons le développement de l'application pour construire de nouvelles fonctionnalités et l'adapter aux retours d'utilisation. L'apport de nouvelles données et l'utilisation des nouveaux outils mathématiques vont nous permettre de progresser dans la modélisation de donnée. Elle deviendra ainsi encore plus pertinente.

SPEC ENVIRONNEMENT