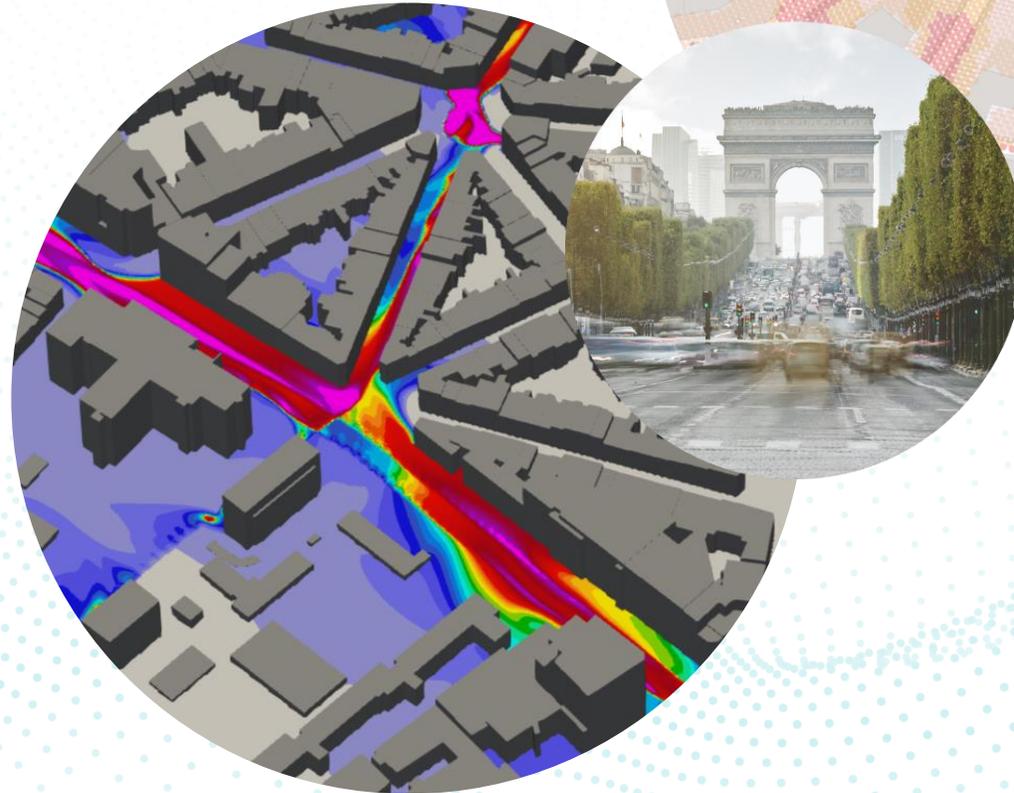


# Webinaire ENGEES

Prise en main des outils classiques de l'IA pour l'environnement – Faisabilité et limites

3 avril 2025



Dr. Xavier Jurado / [xjurado@air-d.fr](mailto:xjurado@air-d.fr)

Dr. Nicolas Reiminger / [nreiminger@air-d.fr](mailto:nreiminger@air-d.fr) / <https://www.reiminger.fr>

Pr. José vazquez / [jose.vazquez@engees.unistra.fr](mailto:jose.vazquez@engees.unistra.fr) / <https://hydraulique-des-reseaux.engees.eu/jose-vazquez>

1. Introduction
2. Généralités
3. Application concrètes de l'IA à l'environnement
4. Quelques points d'attention importants
5. Conclusion et échanges



## **Introduction :**

Présentation des intervenants, mise en contexte et objectifs du webinaire



**Pr. José Vazquez**

Ingénieur ENSAIS

Docteur de l'Université Louis Pasteur de Strasbourg (Sciences de l'eau)

**Professeur à l'ENGEES**

**Chercheur permanent au laboratoire ICUBE (UMR 7357)**



**Dr. Xavier Jurado**

Ingénieur ENGEES

Docteur de l'Université de Strasbourg (IA / SI - Mécanique des fluides)

**Responsable technique AIR&D (C.T.O.) / Chef de projets intelligence artificielle**



**Dr. Nicolas Reiminger**

Ingénieur ENGEES

Docteur de l'Université de Strasbourg (SI - Mécanique des fluides)

**Responsable scientifique AIR&D (C.S.O.) / Chef de projets modélisations numériques**

**Chercheur associé au laboratoire ICUBE (UMR 7357)**

**Membre du comité de la recherche de l'Astee**



### Quelques exemples de défis environnementaux actuels :

- **Changement climatique** : Augmentation des événements climatiques extrêmes (inondations, sécheresses, canicules).
- **Pollution de l'air** : Impact sur la santé publique et les écosystèmes.
- **Gestion des ressources naturelles** : Nécessité d'optimiser l'utilisation des ressources en eau, énergie, etc.
- **Biodiversité en déclin** : Suivi des espèces, des habitats et gestion des zones protégées.



### Limites des approches traditionnelles :

- Modèles physiques **complexes** et **coûteux** à mettre en place.
- Données massives **difficilement exploitables** à grande échelle.
- Besoin de **solutions dynamiques** et **prédictives** pour une gestion intelligente et anticipée .



### Pourquoi l'IA est une réponse :

- **Traitement des données massives** : Exploitation de données provenant de capteurs, satellites, images et simulations.
- **Modélisation prédictive** : permet de prédire des événements rares ou difficiles à simuler directement.
- **Optimisation des décisions** : Aide à la prise de décision pour une gestion plus efficace des ressources.
- **Flexibilité et adaptabilité** : pouvant être capables de s'adapter à des changements rapides et complexes.



### Tendances récentes

- Utilisation croissante des **modèles de Machine Learning** (ex : LSTM, Random Forest) pour les prévisions environnementales.
- **Progrès en IA interprétable** : Aide à comprendre les résultats, ce qui est crucial pour des applications environnementales.
- **Collaboration interdisciplinaire** : IA appliquée à l'environnement avec des sciences sociales, politiques et économiques.





- 1 Acquérir les bases essentielles** pour comprendre et échanger efficacement sur les modèles de Machine Learning et d'Intelligence Artificielle.
- 2 Explorer un panorama des approches IA** : découvrir les grandes familles de techniques et leur pertinence pour des problématiques environnementales.
- 3 Illustrer l'impact concret de l'IA** à travers des études de cas sur l'assainissement, la qualité de l'air, la gestion du risque inondation, etc.
- 4 Comprendre les défis et opportunités** liés à l'application de l'IA (qualité des données, interprétabilité, besoins en calcul, fiabilité des modèles).
- 5 Encourager l'adoption des bonnes pratiques** pour utiliser l'IA de manière efficace, pertinente et responsable dans les projets environnementaux.

# Généralités :

Définitions, principes, différences...

**ALGORITHMIQUE :**

Ensemble de règles opératoires dont l'application permet de résoudre un problème énoncé au moyen d'un nombre fini d'opérations. Un algorithme peut être traduit, grâce à un langage de programmation, en un programme exécutable par un ordinateur.

**INTELLIGENCE ARTIFICIELLE (IA) :**

Ensemble des théories et des techniques développant des programmes informatiques complexes capables de simuler certains traits de l'intelligence humaine

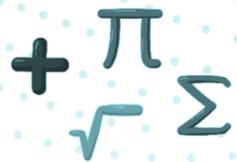
**APPRENTISSAGE AUTOMATIQUE (Machine Learning, ML) :**

Méthodes et techniques permettant d'apprendre des comportements à partir d'un set de données sans règles explicites

**APPRENTISSAGE PROFOND (Deep Learning, DL) :**

Réseaux de neurones composés de plusieurs couches et qui ont la capacité d'apprendre des phénomènes complexes à partir de données

-3000



1950



1980



2010

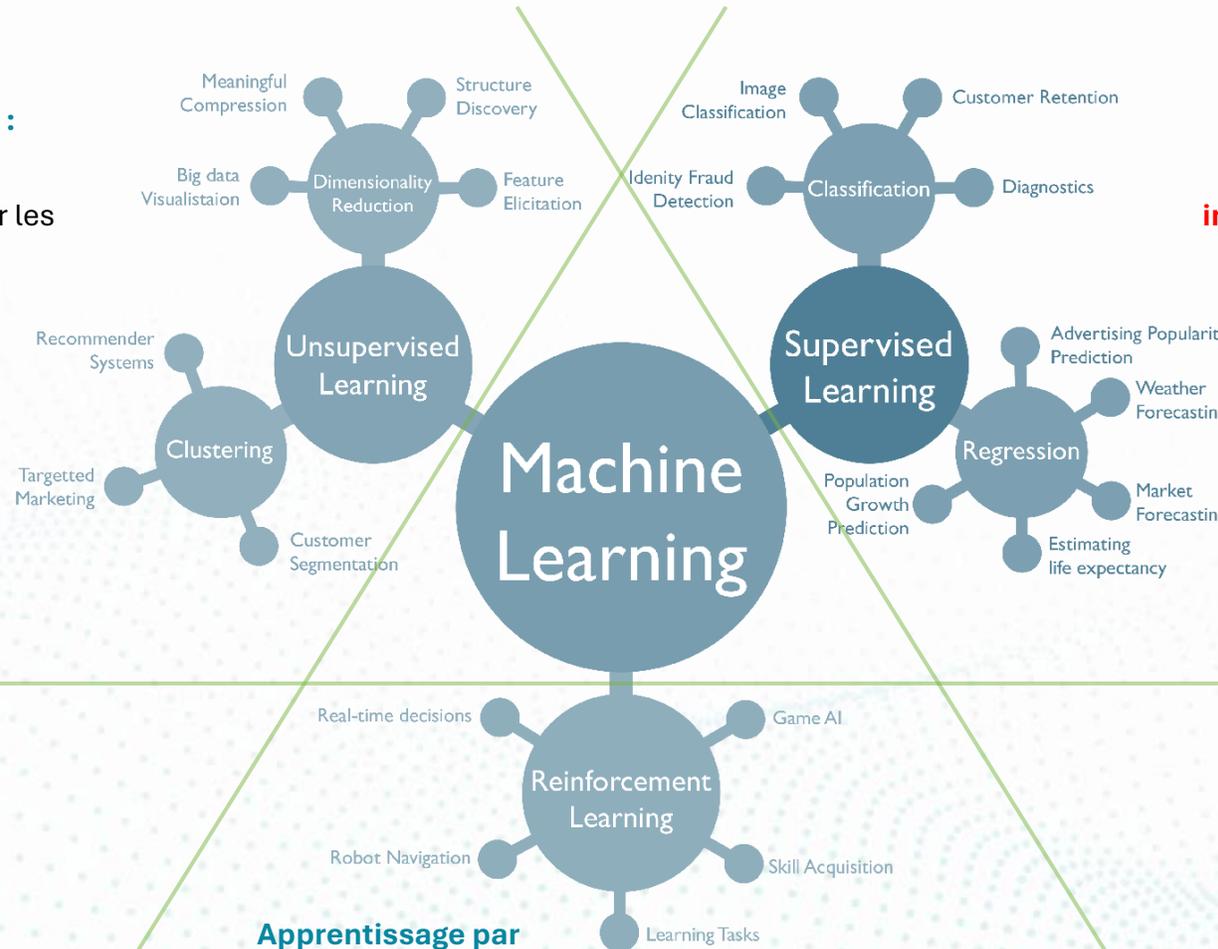


**Apprentissage non supervisé :**

Les données d'entraînement **n'incluent pas** d'informations sur les résultats à atteindre

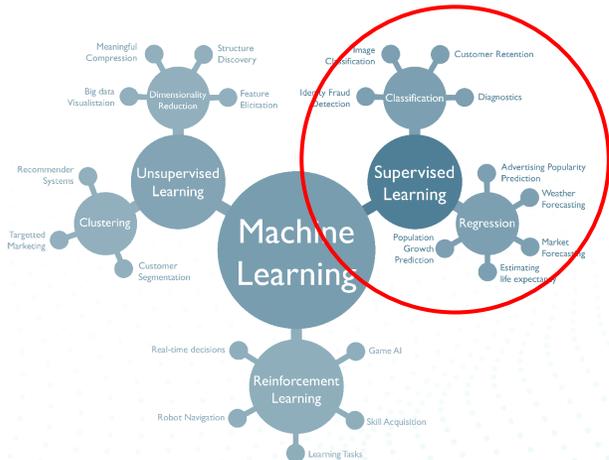
**Apprentissage supervisé :**

Les données d'entraînement **incluent** des informations sur les résultats à atteindre



**Apprentissage par renforcement :**

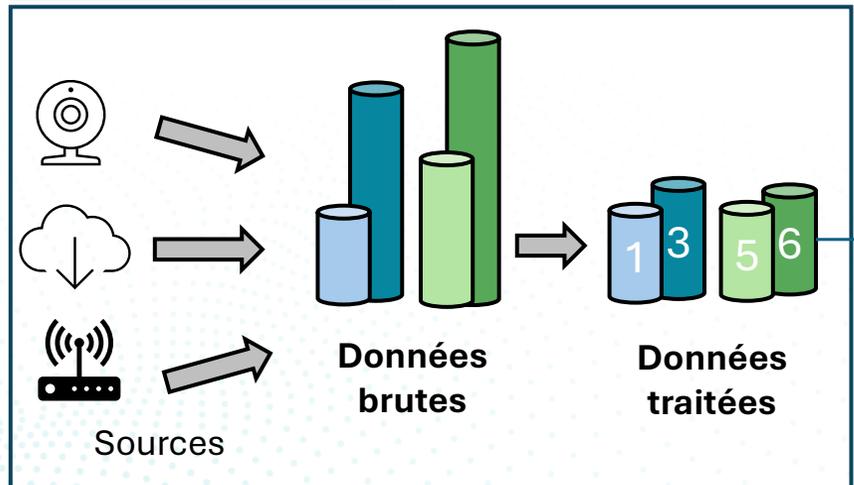
**Environnement évolutif :** les actions modifient l'environnement durant l'entraînement



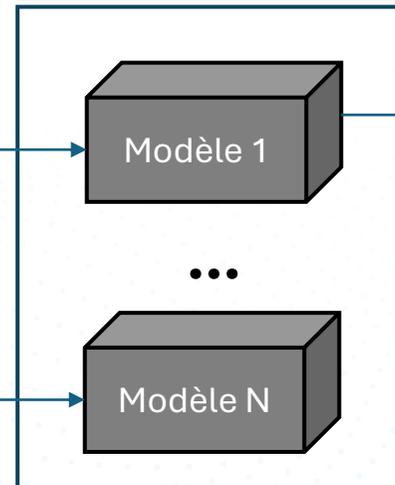
**Apprentissage supervisé :**

Les données d'entraînement **incluent** des informations sur les résultats à atteindre

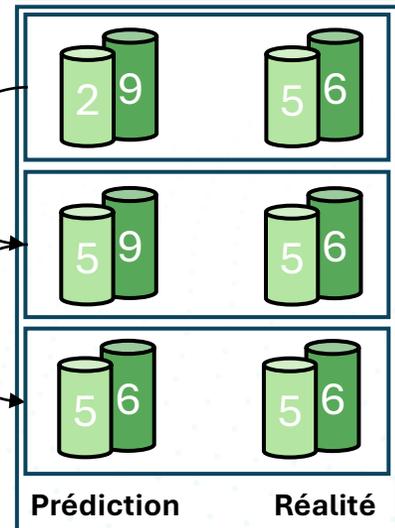
**Pré traitements**



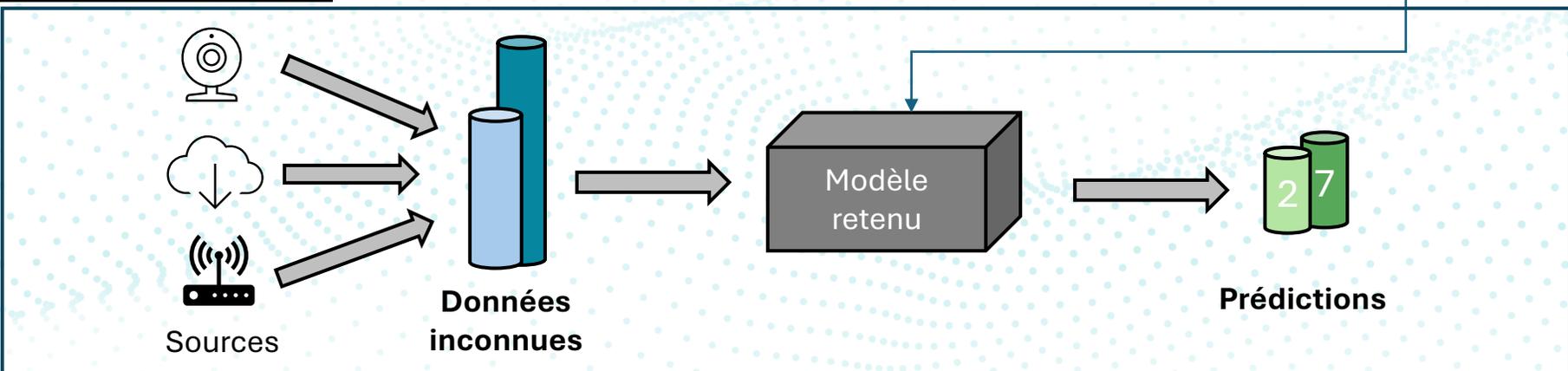
**Choix modèles**

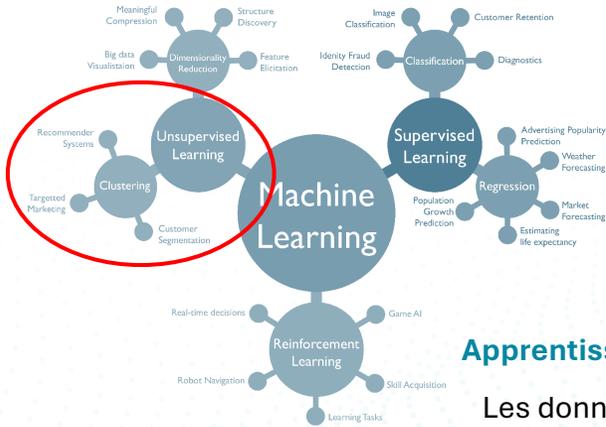


**Entrainement**



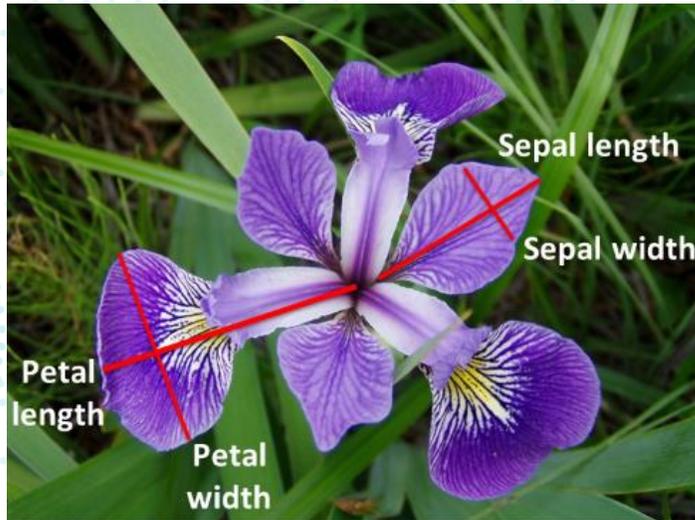
**Utilisation**



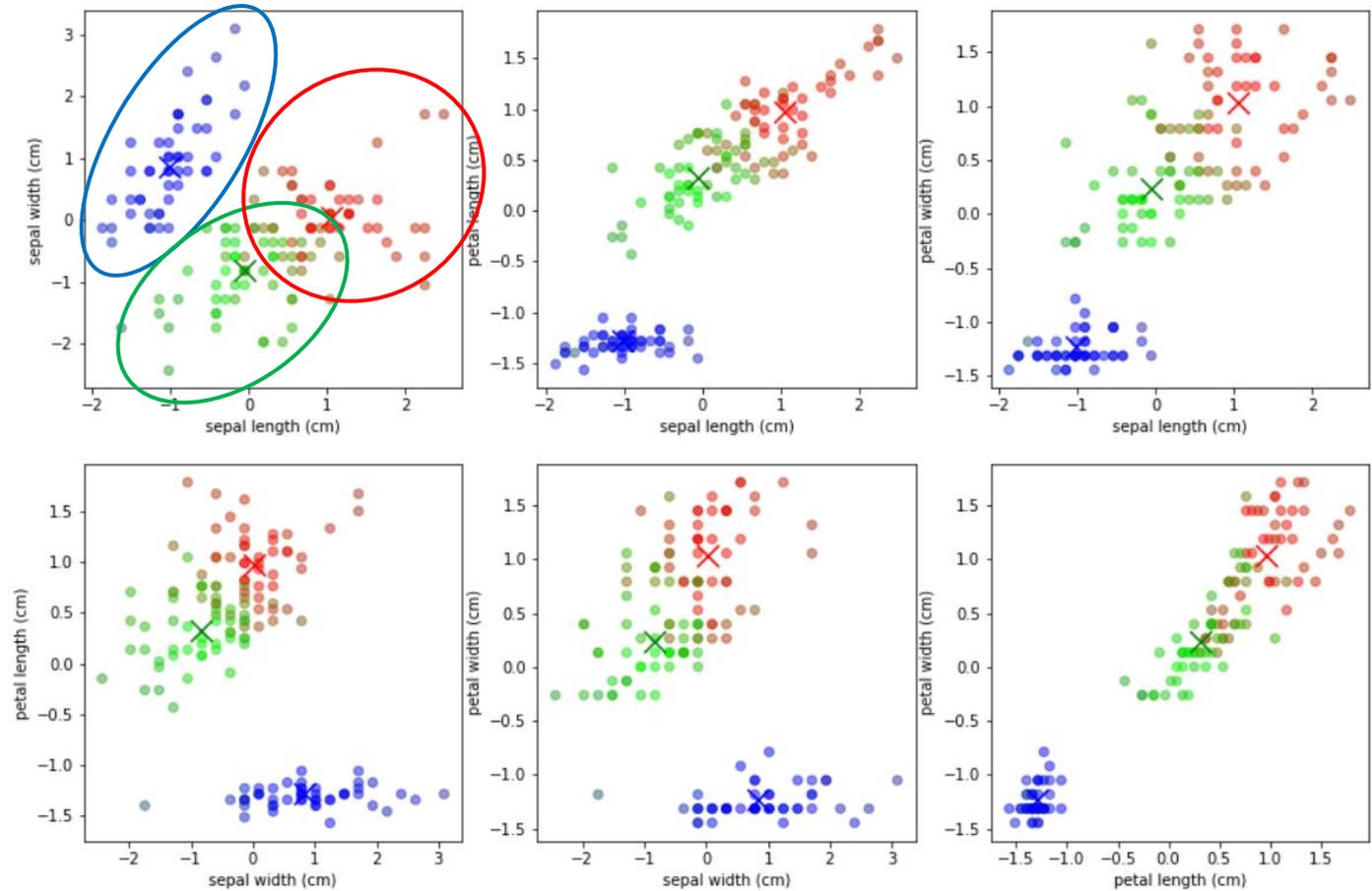


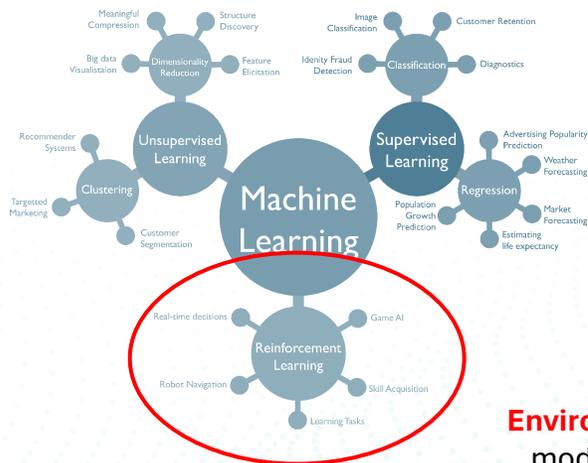
Apprentissage non supervisé :

Les données d'entraînement **n'incluent pas** d'informations sur les résultats à atteindre



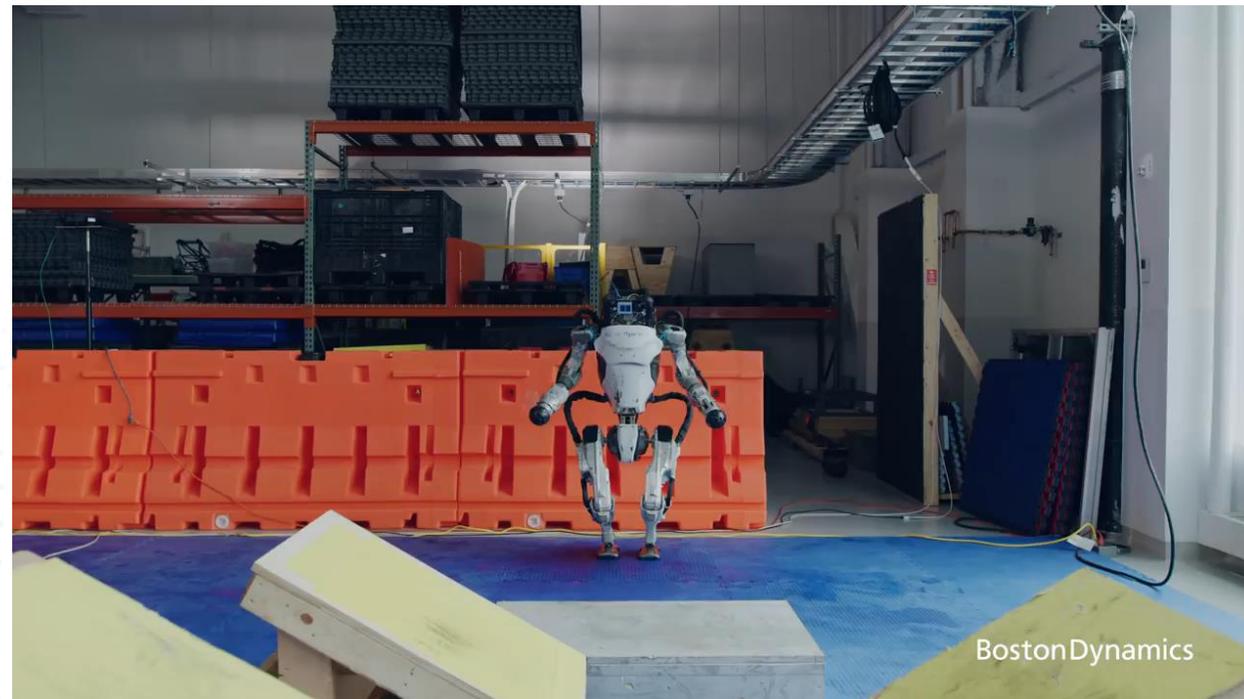
Classification d'espèces végétales (Iris)





### Apprentissage par renforcement :

**Environnement évolutif** : les actions modifient l'environnement durant l'entraînement



BostonDynamics

Robotique

### Voitures autonomes

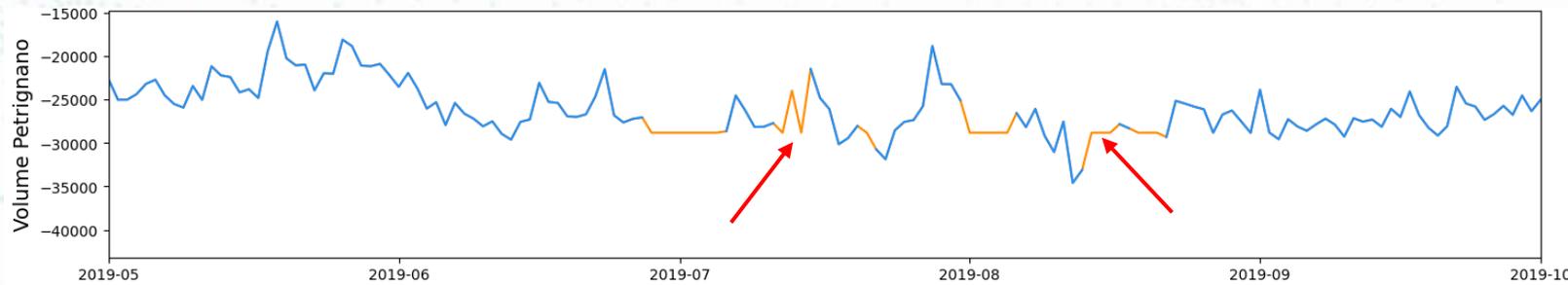
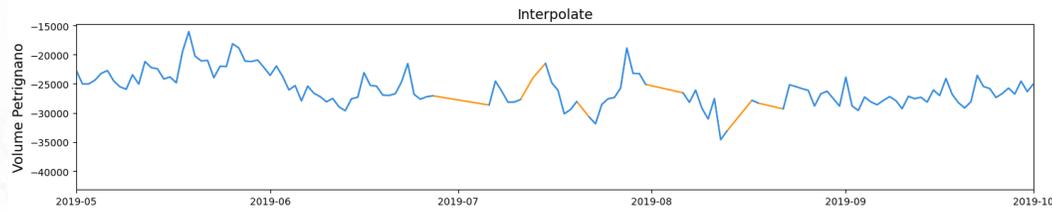
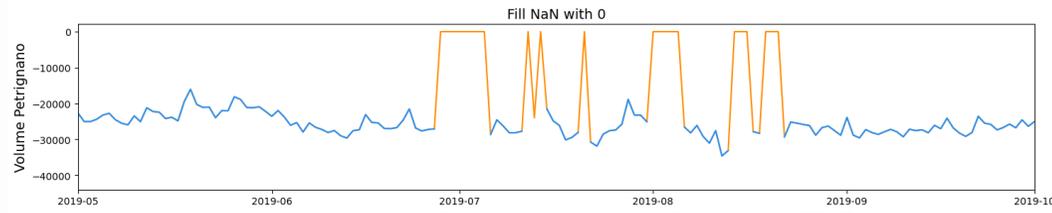


# Applications concrètes de l'IA à l'environnement :

Exemples appliqués à l'assainissement, au milieu  
naturels et la qualité de l'air

## Comment reconstruire des chroniques de données incluant des données manquantes/aberrantes ?

### Cas de la mesure des hauteurs de nappe dans un aquifère

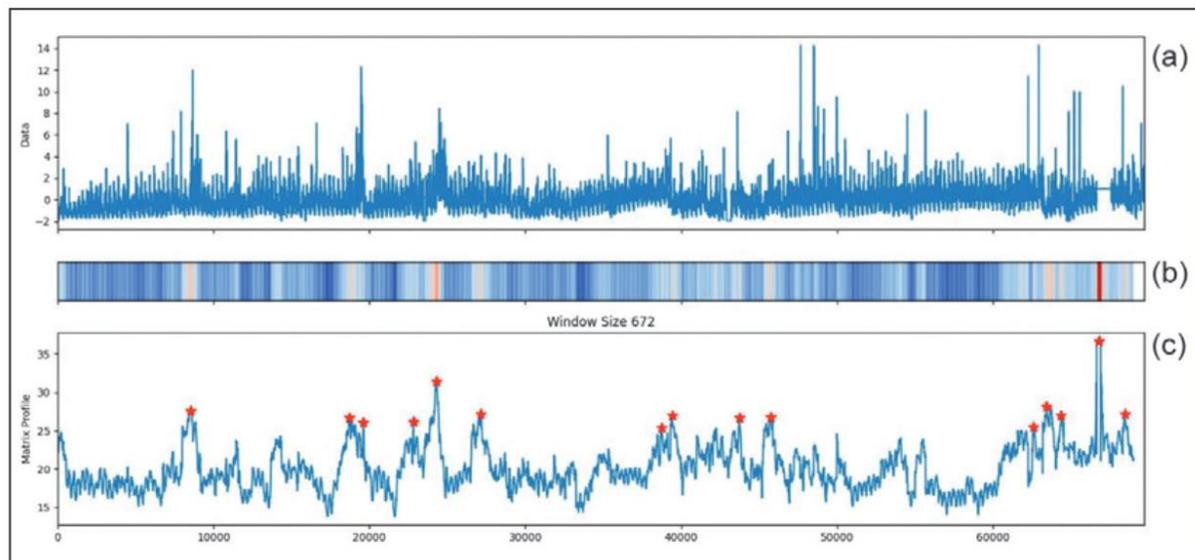


Interpolations permises grâce à des algorithmes « simples » type régression polynomiale, arbres de décision, random forest...

**Apprentissage supervisé (régression)**

**Origine du cas d'illustration :** Formation initiale ingénieur ENGEES

## Comment détecter des anomalies de mesure ?



a) Données brutes de la turbidité normalisée. b) Visualisation du profil matriciel en utilisant une carte de chaleur pour mieux visualiser les anomalies (rouge). c) Valeurs brutes du profil matriciel calculé en fonction du temps. Les étoiles représentent le début des séquences anormales.

Figure 8. Calcul du profil matriciel et identification des anomalies en utilisant Matrix Profile

**Origine du cas d'illustration :** Travaux de thèse Dr. Imane Zidaoui (3DEAU, Pr. José Vazquez) publiés chez TSM, Astee

**Plus d'informations :** Zidaoui, I., Joannis, C., Wertel, J., Isel, S., Wemmert, C., Vazquez, J., & Dufresne, M. (2022). Utilisation de l'intelligence artificielle pour la validation des mesures en continu de la pollution des eaux usées. *Techniques Sciences Méthodes*, 11, 39–51. <https://doi.org/10.36904/tsm/202211039>

## Cas de la mesure de la turbidité en réseau d'assainissement

Détections permises grâce à l'algorithme **Matrix Profile**

Apprentissage non-supervisé

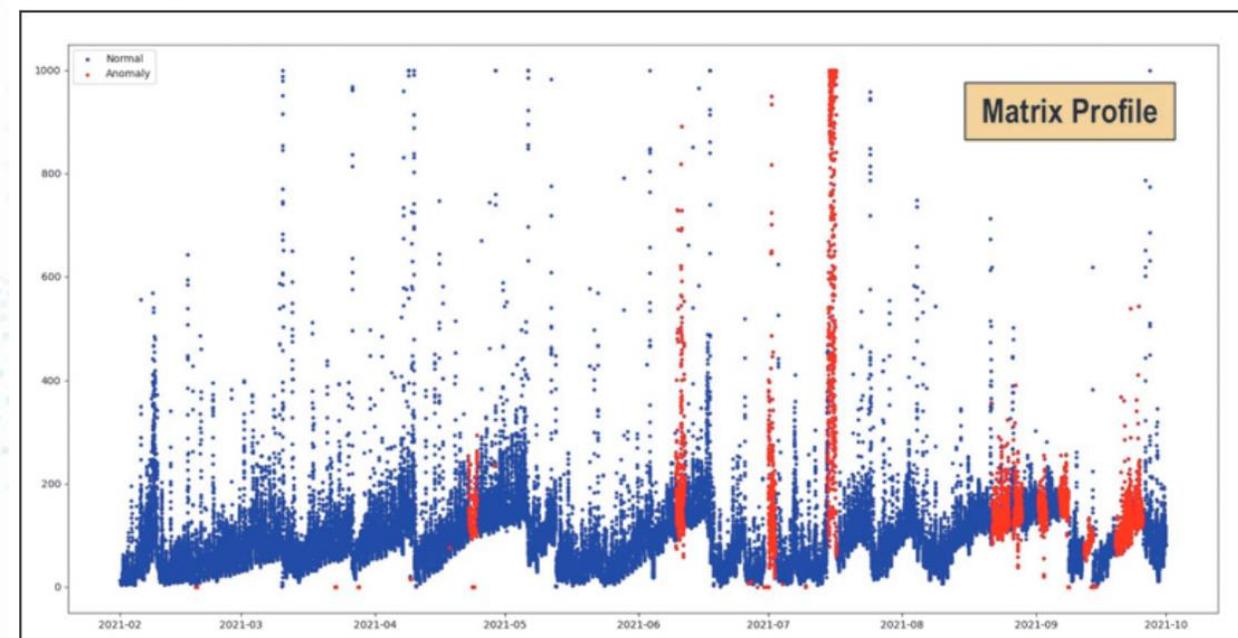


Figure 10. Résultats de la validation des données en utilisant l'algorithme Matrix Profile

## Comment l'IA peut-elle servir à la gestion des risques ?

Détection permise grâce à un algorithme de type convolutifs (CNN) nommé **MobileNet**

**Apprentissage supervisé (classification)**

## Cas de la gestion du risque inondation

Suivi vidéo d'axes routiers/carrefours et analyse des images retournées :



**Pas d'inondation**



**Inondation détectée : alerte**

**Origine du cas d'illustration :** Formation initiale ingénieur ENGEES

## Comment l'IA peut-elle servir à l'écologie ?

Détection permise grâce à des algorithmes de segmentation de type convolutifs (CNN)

Apprentissage supervisé  
(classification)

## Cas du comptage des chauves-souris dans le cadre d'un inventaire faune-flore



Détection et création d'une carte de densité à l'aide de l'algorithme **CSRNet** (Cascaded Selective Resolution Network)

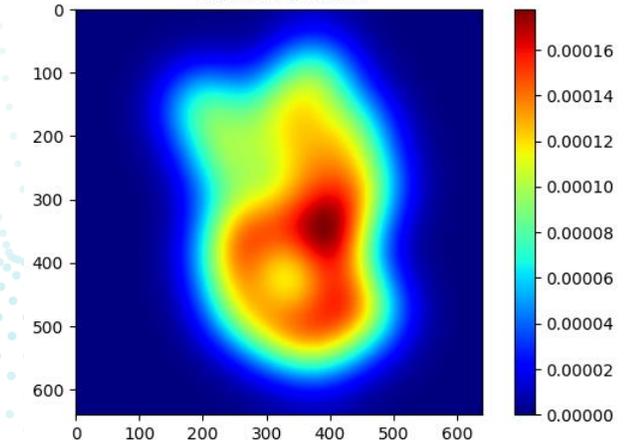
Apprentissage supervisé (classification + traitement des résultats avec interpolation gaussienne)

Origine du cas d'illustration : Travaux de recherche internes AIR&D

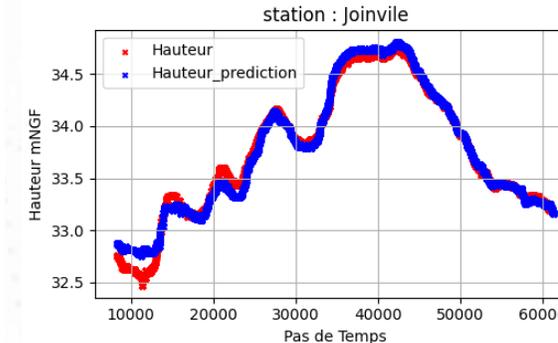
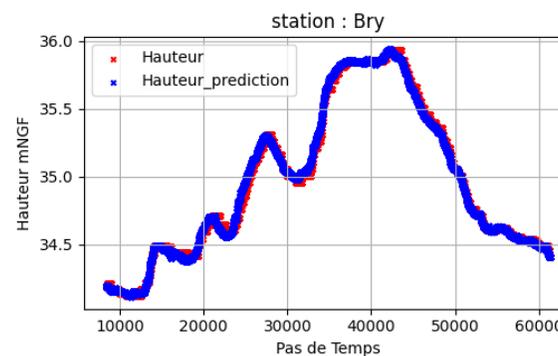
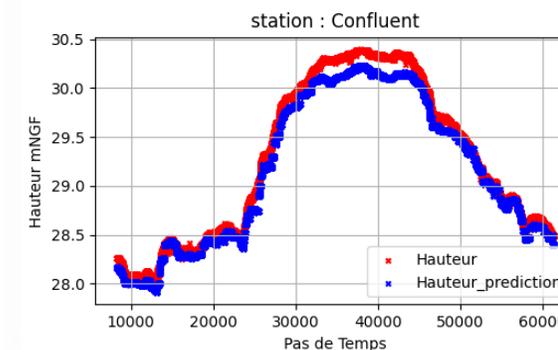
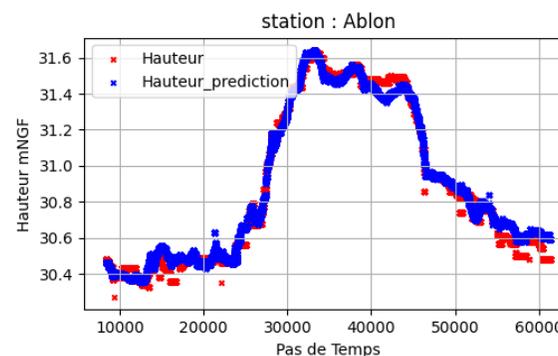
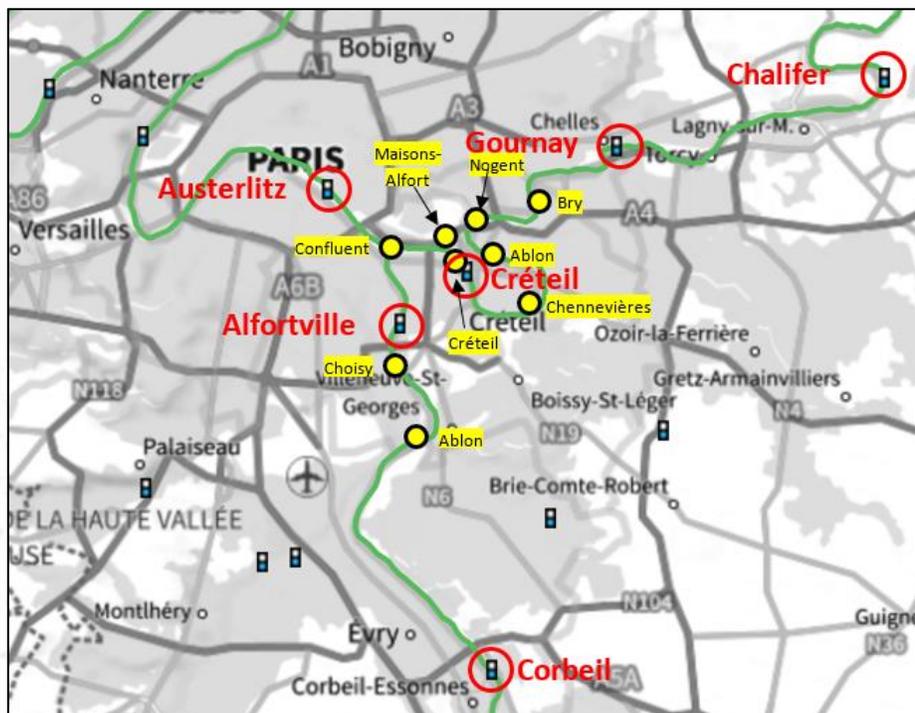
Image d'origine



Carte de Densité



Problème : Comment estimer à l'avance des données issues de capteurs ?



Cas de la prévision des hauteurs d'eau en réseau d'assainissement sur la base des prévision Vigicrue

Prévisions permises grâce à un réseau de neurones artificiels (ANN) type **multiLayerPerceptron**

**Apprentissage supervisé (régression)**

Origine du cas d'illustration : Formation initiale ingénieur ENGEES

Problème : Comment prévoir les crues 24h à l'avance ?

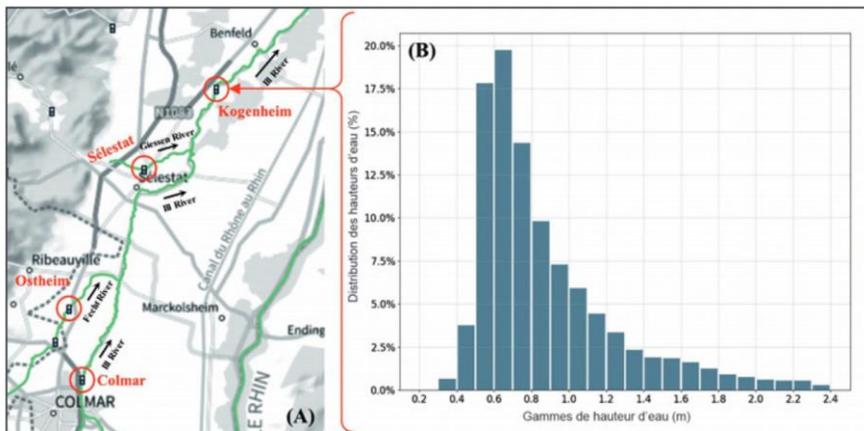
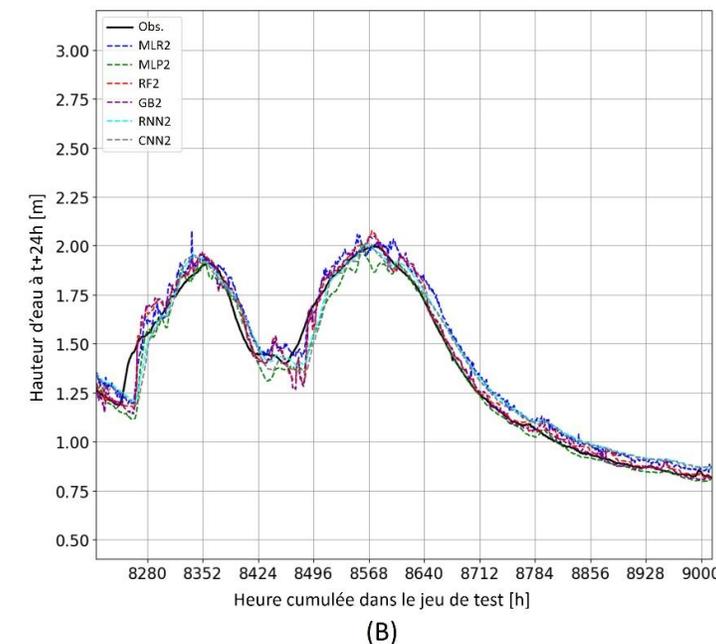
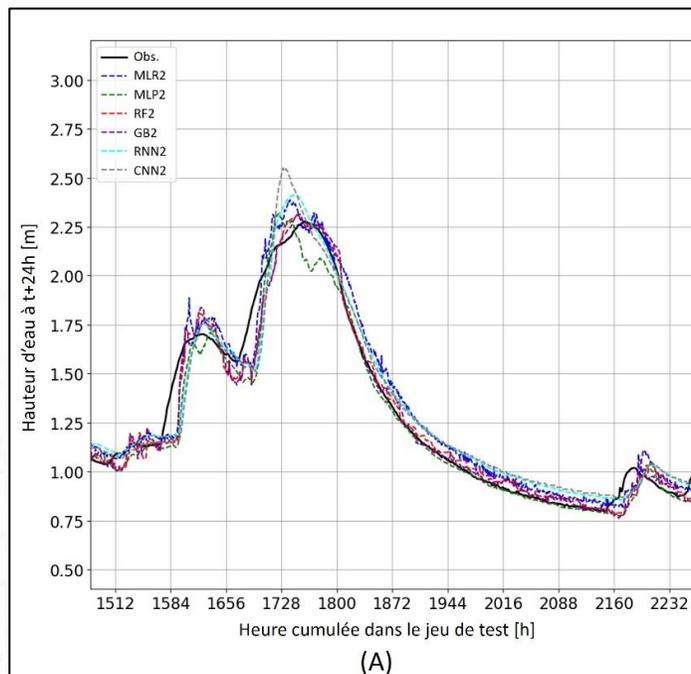
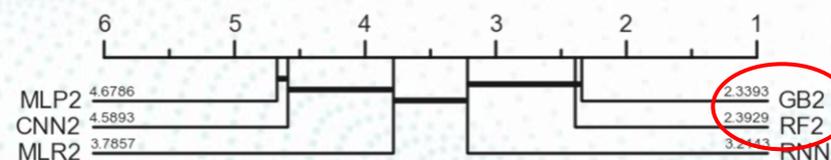


Figure 1. Site principal d'étude. (A) Localisation des stations hydrométriques (www.vigicrues.gouv.fr) et (B) distribution des hauteurs d'eau à Kogenheim entre 2007 et 2022



Meilleurs résultats obtenus à l'aide des algorithmes ensemblistes **Random Forest (RF)** et **Gradient Boosting (GB)**

Apprentissage supervisé (régression)

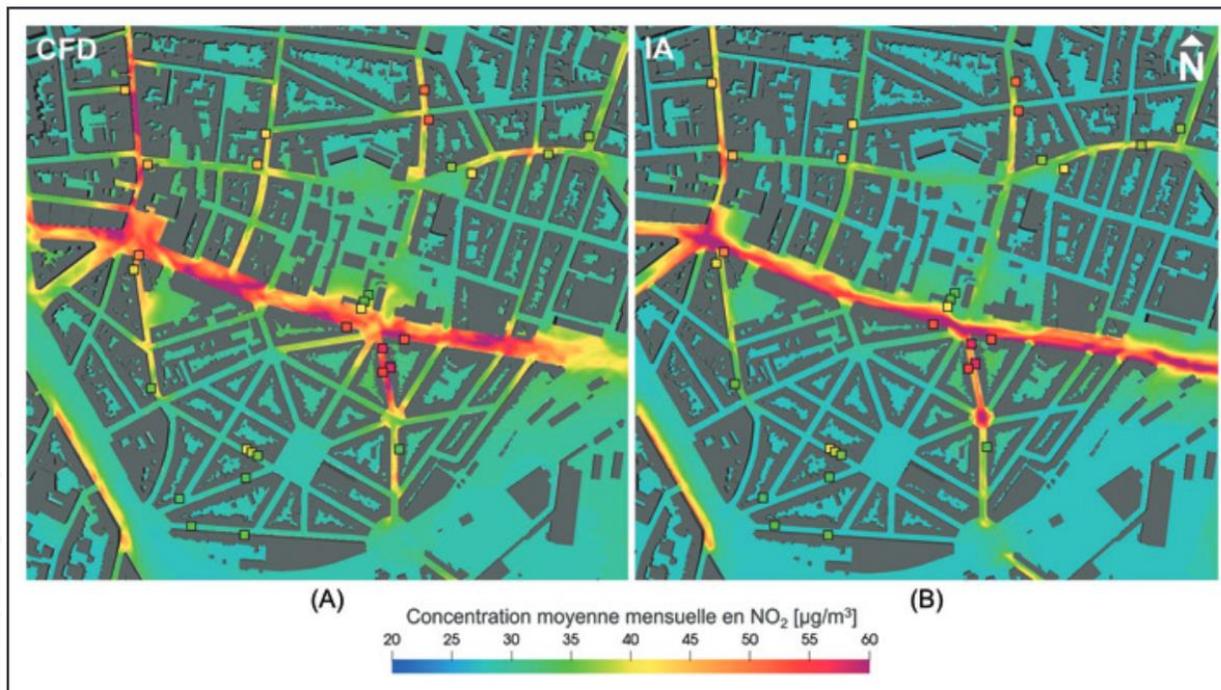


Origine du cas d'illustration : Travaux de recherche internes AIR&D/ICUBE publiés chez TSM, Astee

Plus d'informations : Reiminger, N., Jurado, X., Saunier, L., Maurer, L., Reiminger, E., Weber, L., Ly Nguyen, T. H., & Wemmert, C. (2024). Évaluation des performances de l'intelligence artificielle et de l'apprentissage automatique pour la prévision des crues: Étude de cas du bassin versant de l'Ill. Techniques Sciences Méthodes, TSM 11/2024, 53–71.

<https://doi.org/10.36904/tsm/20241153>

**Problème : comment modéliser la qualité de l'air en ville avec précision et rapidité ?**



**Figure 5. Cartographie des concentrations moyennes mensuelles en dioxyde d'azote NO<sub>2</sub> modélisées (A) avec le modèle CFD et (B) avec le modèle IA pour la période du 30 avril au 28 mai 2016 et comparaison avec les concentrations moyennes relevées sur site à l'aide des analyseurs passifs**

Modèle de **réseaux de neurones convolutifs** (CNN) type « encoder-decoder »

Apprentissage supervisé (régression)

Modèle	FAC2	MFE	NMSE	MRE	R	NSD	Target
CFD	1,00	0,14	0,02	0,13	0,76	0,89	0,68
IA	1,00	0,17	0,06	0,18	0,72	0,82	1,20
Objectif	> 0,5*	–	< 1,5*	< 0,3***	> 0,65**	–	< 0,80**
Valeur parfaite	1	0	0	0	1	1	0

\*D'après CHANG et HANNA [2005], \*\*d'après THUNIS *et al.* [2012], \*\*\*d'après EU [2008] pour la modélisation annuelle des concentrations en NO<sub>2</sub>.

Tableau II. Comparaison entre modèles CFD et IA sur les résultats de concentrations mensuelles en NO<sub>2</sub> pour les différents critères de performances considérés

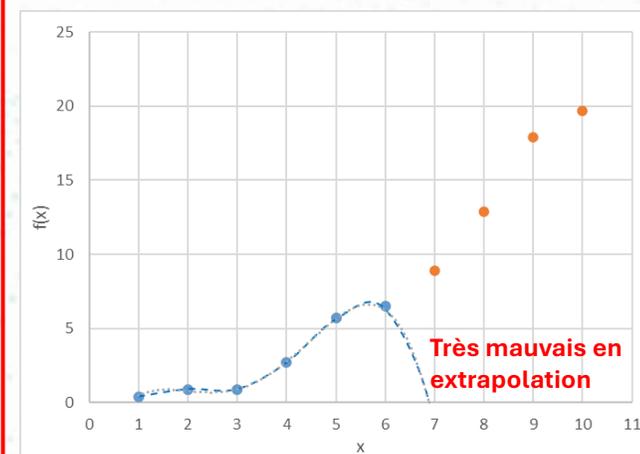
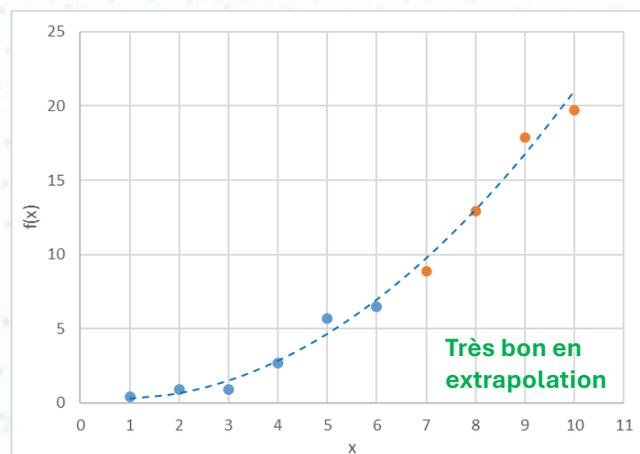
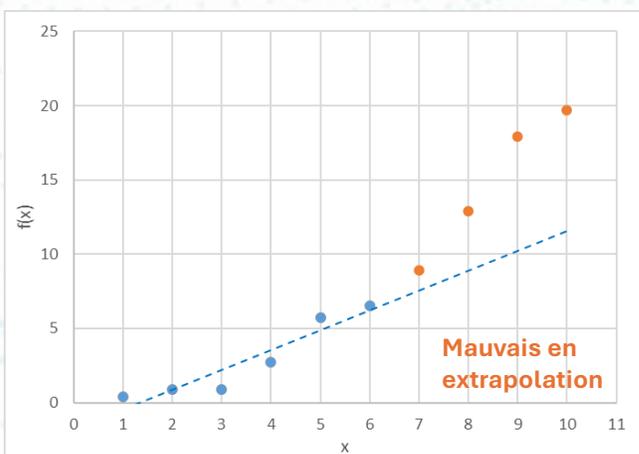
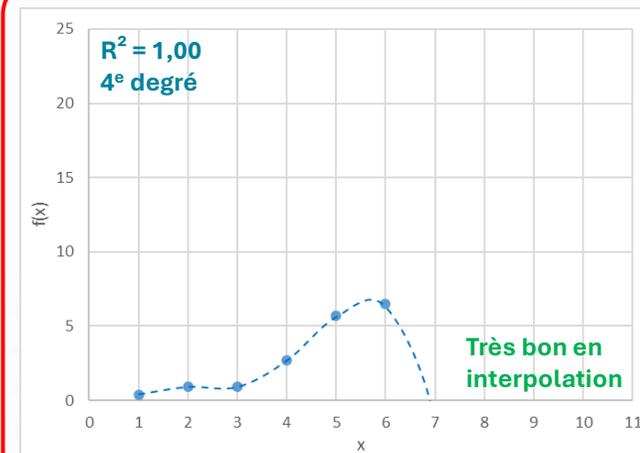
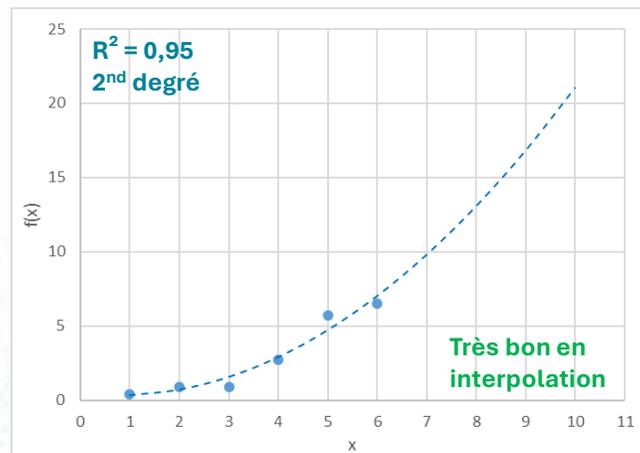
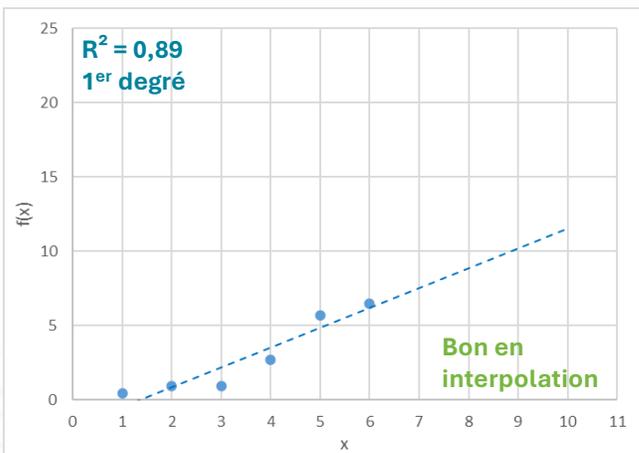
**Origine du cas d'illustration :** Travaux de recherche internes AIR&D/ICUBE publiés chez TSM, Astee

**Plus d'informations :** Reiminger, N., & Jurado, X. (2024). Modélisation micro-échelle de la qualité de l'air en milieu urbain par méthodes CFD et IA. *Techniques Sciences Méthodes*, 9, 41–53. <https://doi.org/10.36904/20240941>

# Quelques points d'attention importants :

Surapprentissage, qualité des données, interprétabilité...

### Exemple des régressions polynomiales



### Exemple typique de surapprentissage

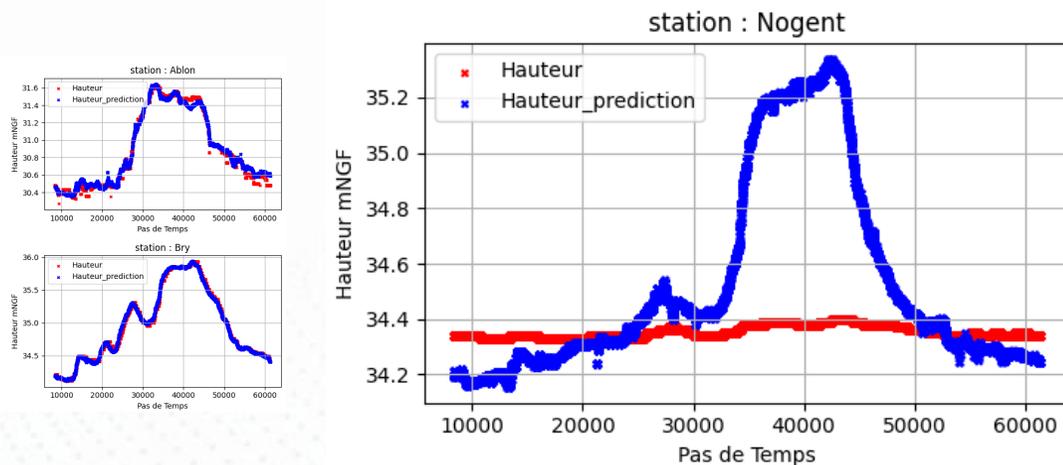
Le degré de liberté de l'équation (degré du polynôme) est devenu si grand que l'algorithme a « appris par cœur » pour parfaitement suivre les données d'entraînement.

#### Conséquences :

- La généralisation du modèle et son application sur des exemples extrêmes, en dehors du champ des données d'entraînement (extrapolation) devient **mauvaise**
- L'applicabilité du modèle sur des données comprises dans le champ des données d'entraînement (interpolation) peut également devenir **mauvaise** (non visible sur cet exemple simple)

### Exemple de résultat avec des données erronées

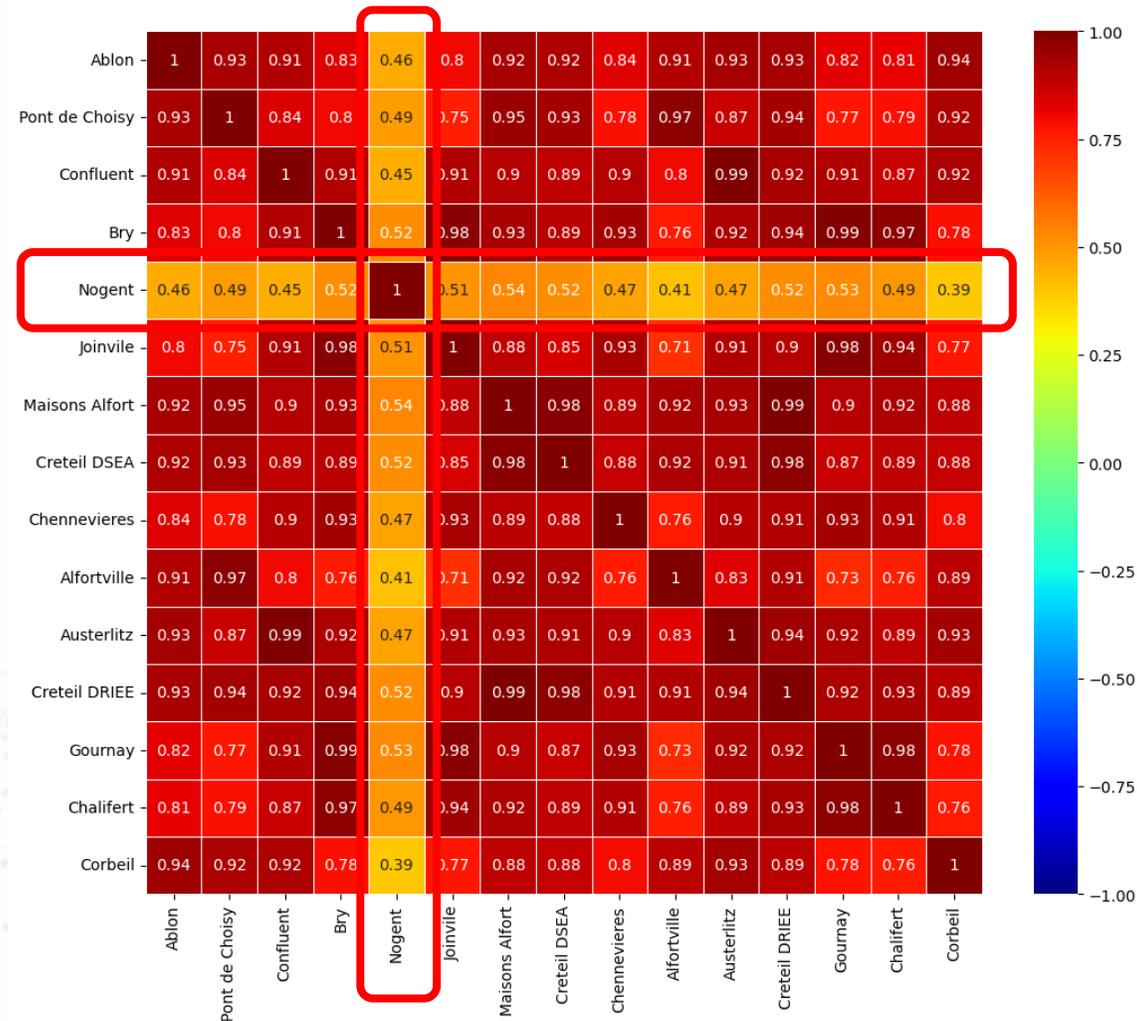
Cas de la prévision des hauteurs d'eau en réseau d'assainissement



La qualité des résultats issus d'un modèle de machine learning/d'IA (plus généralement, d'un modèle numérique) **dépend directement de la qualité de la donnée d'entrée.**

Il est toujours utile, en tout premier lieu, d'**étudier la donnée d'entrée** (matrice de corrélation, diagramme pairplot, etc.) pour anticiper tout problème !

Matrice de corrélation

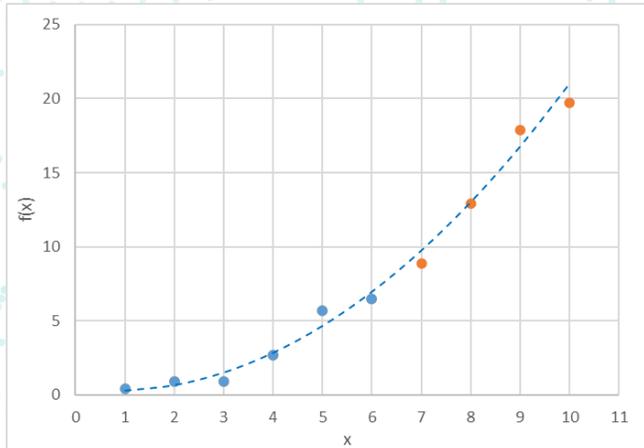
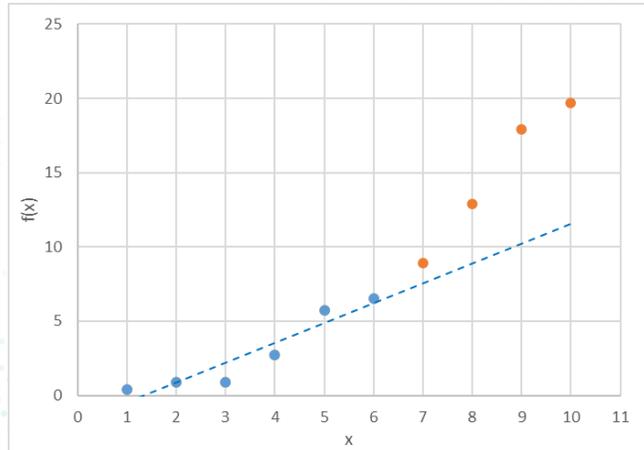


Matrice de corrélation des stations disponibles

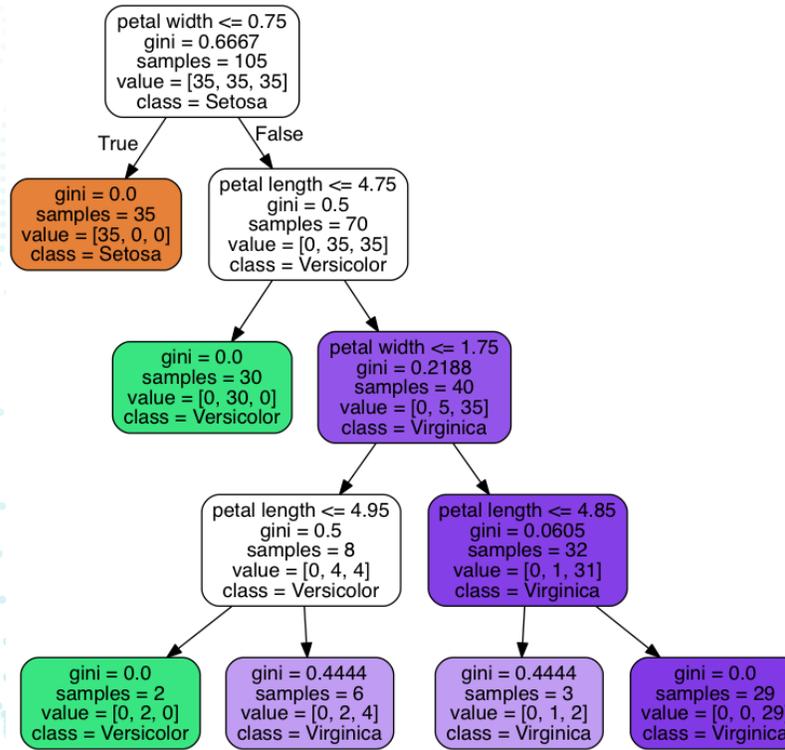
(1 : corrélé positivement, -1 corrélé négativement, 0 absence de corrélation)

L'interprétabilité des modèles dépend directement de leur complexité

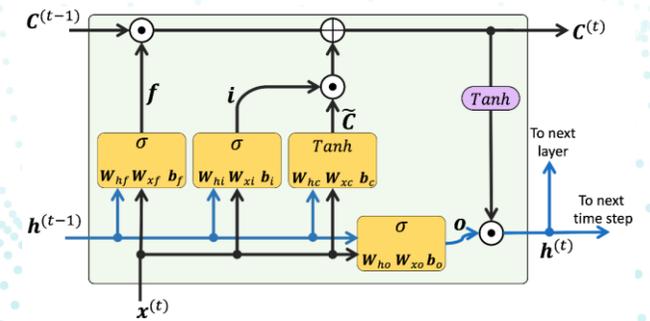
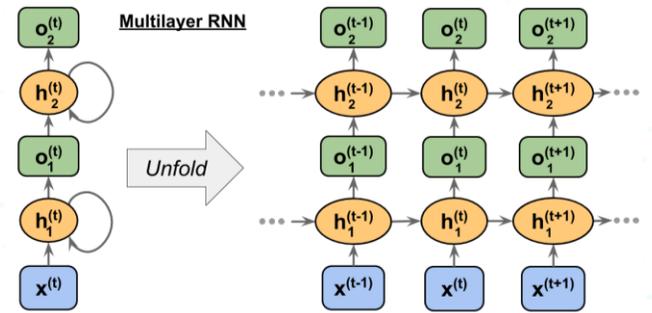
Exemple n°1 : évolution polynomiale (facile)



Exemple n°2 : arbre de décision (interprétable, mais rapidement complexe)



Exemple n°3 : réseaux de neurones (pas interprétable directement)



Illustrations : Raschka, Sebastian, and Vahid Mirjalili. Python Machine Learning, 3rd Ed. Packt Publishing, 2019.

# Conclusion et échanges :

## Synthèse et questions/réponses

## Bilan :

- **L'IA est un outil puissant** pour traiter des problématiques environnementales complexes (ressource en eau, hydrologie, climat, biodiversité, qualité de l'air, etc).
- **Différentes approches existent** : IA symbolique, Machine Learning, Deep Learning, chacune ayant ses forces et limites.
- **L'IA pour l'environnement repose sur des données** : leur quantité, qualité et leur accessibilité sont des défis majeurs, **mais aussi sur des bonnes pratiques**.
- **L'interprétabilité des modèles** est essentielle pour une adoption large et une prise de décision éclairée, mais reste encore difficile aujourd'hui.
- **Les défis restent nombreux** : contraintes de calcul, robustesse des modèles, confiance et acceptabilité par les acteurs.
- **Les perspectives sont prometteuses** avec l'évolution des capacités de calcul, la démocratisation des outils et l'amélioration des algorithmes.

## Enjeux actuels :

- Intégrer l'IA dans des **systèmes de décision** concrets
- Développer des modèles plus **transparents** et **explicables**
- Associer IA et modélisation physique pour une **approche hybride**
- Travailler en **collaboration entre ingénieurs, chercheurs et décideurs**

## Questions, réponses

