

Prédiction de la porosité dans des réservoirs à hydrogène haute pression en matériaux composites par IA Cognitive Floue Augmentée en lien avec les paramètres du procédé de mise en œuvre

Prediction of Porosity in High-Pressure Hydrogen Tanks Made of Composite Materials Using Augmented Fuzzy Cognitive AI in Relation to Manufacturing Process Parameters

Lina ACHOUR^{1,2}, Zyed ZALILA^{3,4}, Zoheir ABOURA¹, Benjamin LORENTZ⁴, David RUGGI², Kamel KHELLIL¹

1 : Roberval
Université de Technologie de Compiègne
Royallieu Research Center, CS 60319-60203, Compiègne
e-mail : lina.achour@utc.fr, zoheir.aboura@utc.fr et kamel.khellil@utc.fr

2 : H2-Power
OPmobility
ZAC du Bois de Plaisance, 214 Avenue de la Mare Gessart, 60280 Venette
e-mail : david.ruggi@opmobility.com

3 : Costech, EA 2223
Université de Technologie de Compiègne, 60200 Compiègne

4 : IDAGRAI LABS / Intellitech
60610 Lacroix Saint-Ouen
e-mail : zyed.zalila@intellitech.fr et benjamin.lorentz@intellitech.fr

Résumé

Les réservoirs d'hydrogène haute pression de type IV (HPV) sont essentiels pour les applications industrielles et les véhicules à hydrogène. Fabriqués à partir de matériaux composites, notamment des fibres de carbone, ces réservoirs sont légers et résistants à la fatigue. Toutefois, des défauts internes comme les porosités peuvent compromettre leur intégrité. Bien que certaines porosités soient tolérées, une présence excessive peut entraîner des défaillances mécaniques. Cette étude vise à développer un modèle prédictif du nombre de porosités en fonction des paramètres du procédé de fabrication. Pour surmonter les limites des approches physiques, nous utilisons l'Intelligence Artificielle Raisonnée Générale (IARG) Xtractis, qui permet la découverte automatique de modèles robustes et intelligibles. À partir de 58 prédicteurs potentiels, 15 ont été sélectionnés pour créer un modèle prédictif. Les résultats montrent une performance réelle élevée avec un RMSE de 7,94 % et une corrélation de 0,824 sur un jeu de test externe. L'intelligibilité des modèles générés par Xtractis est un avantage majeur par rapport à d'autres techniques non-linéaires comme Boosted Tree, qui manquent de transparence. L'approche Xtractis offre ainsi de nouvelles perspectives pour améliorer la gestion de la qualité des réservoirs.

Abstract

High-pressure hydrogen storage vessels of type IV (HPV) are essential for industrial applications and hydrogen vehicles. Made from composite materials, particularly carbon fibers, these vessels are lightweight and fatigue-resistant. However, internal defects such as porosities can compromise their integrity. While some porosity is tolerated by certification standards, excessive presence can lead to mechanical failures. This study aims to develop a predictive model for the number of porosities based on manufacturing process parameters. To overcome the limitations of physical approaches, we use the Xtractis General Reasoning Artificial Intelligence (GRAI), which enables the automatic discovery of robust and intelligible models. Out of 58 potential predictors, 15 were selected to create a predictive model. The results show high performance with a 7.94% RMSE and a correlation of 0.824 on an external test set. The intelligibility of models generated by Xtractis is a major advantage compared to other nonlinear techniques like Boosted Tree, which lack transparency. The Xtractis approach offers new perspectives for improving the quality management of vessels.

Mots Clés : Porosité, réservoirs d'hydrogène haute pression (HPV), tomographie par rayons X, Xtractis / IA Cognitive Floue

Keywords : Porosity, high-pressure hydrogen vessels (HPV), X-ray tomography, Xtractis / Fuzzy Cognitive AI

1. Introduction

Les réservoirs d'hydrogène haute pression de type IV (HPV) sont essentiels dans les systèmes de stockage et de transport d'hydrogène, notamment pour les applications industrielles et les véhicules à hydrogène. Conçus à partir de matériaux composites, principalement avec des fibres de carbone, ces réservoirs intègrent un liner en polymère thermoplastique. Leur structure légère et robuste, grâce à l'utilisation de composites, leur confère une grande résistance à la fatigue, ce qui prolonge leur durée de vie tout en préservant leur intégrité mécanique [1]. De plus, leur faible densité contribue à la réduction de la masse globale tout en optimisant la capacité de stockage. Cependant, la présence de défauts internes, tels que les porosités, peut compromettre l'intégrité structurelle de ces réservoirs [2] [3]. Bien qu'une certaine porosité soit tolérée dans les normes de certification, il est crucial de pouvoir détecter et contrôler ces défauts afin de garantir la sécurité des réservoirs. Des porosités excessives ou mal localisées peuvent entraîner des défaillances mécaniques [4] [5]. Ces porosités sont étroitement liées à de nombreux paramètres du procédé de fabrication, en particulier lors de l'enroulement filamentaire. En effet, il existe 56 paramètres potentiellement contrôlables pendant la production, tels que la tension des bobines de fil, la température de la résine, le doctor blade (qui contrôle la quantité de résine déposée), et l'angle d'enroulement, pour n'en citer que quelques-uns. Il est toutefois illusoire de tenter de lier ces paramètres à la taille et au nombre de porosités à travers des approches physiques. Dans ce contexte, il semble pertinent d'adopter une approche basée sur le traitement de données pour développer des modèles prédictifs du nombre de porosités susceptibles d'apparaître dans les réservoirs en fonction des paramètres du procédé. L'objectif étant de pouvoir ajuster ces paramètres pour minimiser la formation de porosités et en améliorer ainsi le contrôle. Une approche utilisant des réseaux de neurones a été écartée en raison du manque de transparence des résultats obtenus. Nous proposons plutôt d'explorer des solutions avancées basées sur l'IA Cognitive Xtractis, une technologie développée par la société INTELLITECH.

L'IA Raisonnante Générale (IARG) XTRACTIS propose une approche disruptive de modélisation prédictive : elle permet la découverte automatique de modèles à la fois robustes et intelligibles, c'est-à-dire performants sur des cas inconnus et totalement transparents. Les modèles induits sont composés de règles SI...ALORS à logiques continues ; chaque règle est définie par une Relation Floue d'ordre N qui est en fait une équation non-linéaire reliant la variable à prédire aux N-1 prédicteurs retenus par XTRACTIS [6].

L'IARG instancie les trois modes de raisonnement humain :

- **L'induction** pour la découverte de modèles prédictifs robustes et intelligibles à partir d'observations
- La **déduction** pour la prédiction à très haute fréquence à partir des modèles induits
- **L'abduction** pour la découverte des solutions les plus optimales satisfaisant une requête multiobjectif non-linéaire sous contraintes flexibles, construite à partir de modèles induits ou de modèles analytiques.

L'induction humaine utilisée en sciences expérimentales depuis le 11^e siècle consiste à produire, grâce à la réflexion humaine et à partir d'observations entrées-sortie, une équation ou un système d'équations modélisant le processus/phénomène complexe (PPC) étudié. Ce modèle analytique doit non seulement fonctionner pour les cas connus, ceux qui ont permis de l'obtenir, mais surtout pour

les cas inconnus. À la différence de l'approche cartésienne en sciences expérimentales (induction humaine réductionniste, causaliste, incrémentale binaire (Descartes 1637)), l'induction XTRACTIS est automatique holistique décrémente floue [7]; elle autorise la modélisation de PPC grâce à la prise en compte des signaux faibles et des interactions contributives entre prédicteurs, ce que ne permet nullement l'induction cartésienne.

2. Méthodologie

Pour notre étude, 58 prédicteurs potentiels étaient disponibles (paramètres du procédé de mise en œuvre), représentant des paramètres de conception et de processus critiques qui influencent la qualité et la performance des réservoirs. Ces variables incluent des facteurs liés à la production, tels que la tension appliquée durant l'enroulement, la vitesse de dépose des couches, ainsi que des informations détaillées sur le type d'enroulement utilisé (couches circulaires ou hélicoïdales). Ces paramètres de conception et de processus sont essentiels pour comprendre la relation entre les conditions de fabrication et l'apparition de défauts, notamment les porosités, au sein des matériaux composites. L'un des capteurs retournait la valeur zéro chaque fois qu'il ne pouvait mesurer. Ces valeurs ont donc été remplacées par le label « *missing value* » afin d'éviter un fort biais dans le calcul de prédicteurs caractéristiques de chaque réservoir. Les robots XTRACTIS interprètent « *missing value* » comme le fait que toutes les valeurs de l'univers du discours sont tout à fait possibles.

Deux difficultés majeures ont été rencontrées. La première est le faible nombre de cas de référence et leur distribution : nous ne disposons que de 12 réservoirs composés de 15 couches, soit 180 cas en tout. La deuxième difficulté concerne le traitement des données effectué en amont : les mesures réalisées sur les réservoirs constituaient des signaux temporels complexes dont seules les valeurs minimum, maximum et moyenne par couche d'enroulement ont été gardées. D'autres informations auraient pu être extraites du signal d'origine pour enrichir les données et par conséquent améliorer les chances de découverte de modèles robustes. Cela constitue un avantage notable par rapport aux méthodes classiques de traitement de données.

Les cas de la base de référence sont placés aléatoirement dans différents sous-ensembles TrD, VD, ou ETD.

- **Training Dataset (TrD)** : points de référence connus du robot et utilisés pour créer le modèle et évaluer sa performance **descriptive**
- **Validation Dataset (VD)** : points de référence **inconnus** du robot et utilisés pour évaluer la performance **prédictive** du modèle créé à partir de TrD
- **External Test Dataset (ETD)** : points de référence **inconnus** du robot et utilisés pour évaluer la performance **réelle** du modèle créé à partir de TrD et pour comparer les performances réelles des modèles créés par les techniques IA en compétition.

Les 180 points de référence ont été séparés en 153 points pour l'ensemble $\text{TrD} \cup \text{VD}$ et 27 points pour ETD.

3. Résultats obtenus

Après analyse de la base et en raison de la distribution des valeurs en sortie, il a été choisi de modéliser le nombre de porosités en \log_{10} .

Le processus de modélisation a consisté à induire, à partir des 153 points d'apprentissage, un Collège d'Experts Virtuel (CVE2) constitués de 100 modèles unitaires XTRACTIS appelés Experts Virtuels Individuels (IVE) dont les décisions sont agrégées. 2 000 stratégies d'induction (is) ont été

explorées, de manière compétitive et évolutive, et avec à chaque fois 3 opérateurs d’agrégation (ao). Le meilleur CVE2 retenu (évalué par validation croisée à n splits × k-folds) a ensuite été utilisé pour générer une base de données synthétiques d’environ 20 000 points à partir de ses propres prédictions dans le voisinage de ses points d’apprentissage. À partir de cette nouvelle base étendue, XTRACTIS induit finalement un unique IVE intelligible, copiant la stratégie de décision robuste du CVE, et dont les performances peuvent finalement être évaluées grâce à l’ETD. Cette opération est appelée Reverse Engineering du CVE2 (REV) ; elle permet d’induire des modèles prédictifs robustes et intelligibles lorsque le nombre de cas de référence est faible.

Les modèles XTRACTIS sont comparés à des modèles obtenus par Boosted Tree.

Pour la variable à prédire **log10(Nombre de porosités)**, les performances du **CVE2 XTRACTIS** sélectionné sur le RMSE en validation après un 20 × 5 folds sont les suivantes, comparées au hasard :

log10 (nb of porosities)	Correlation	MAE	RMSE	Refusal
Descriptive Performance / Training (2,000is × 3ao. 20sp × 5f × 80%)	0.943	0.13 (4.15%)	0.18 (5.80%)	0 (0.00%)
Predictive Performance / Validation (2,000is × 3ao. 20sp × 5f × 20%)	0.804	0.21 (6.77%)	0.30 (9.69%)	0 (0.00%)
Real Performance / External Test	0.830	0.20 (6.37%)	0.24 (7.63%)	0 (0.00%)
Performance against chance to beat / External Test P-Value = 0.001 (among 100,000 random permutations)	<i>0.585</i>	<i>0.28 (8.93%)</i>	<i>0.37 (11.89%)</i>	

Tab. 1. Performances du CVE2 XTRACTIS : log10(Nombre de porosités)

MAE = Mean Absolute Error et RMSE = Root Mean Square Error

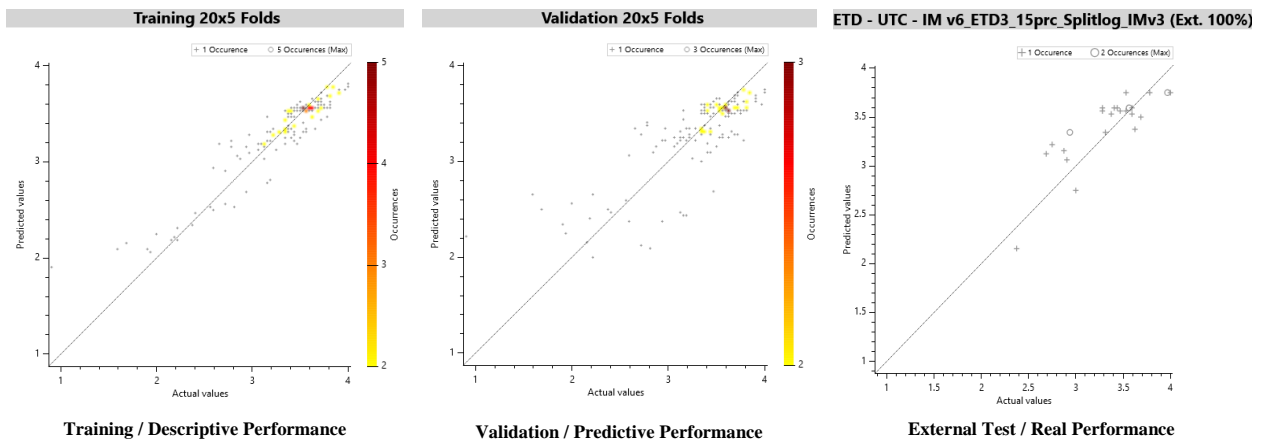


Fig. 1. Graphes de performances du CVE2 XTRACTIS : log10(Nombre de porosités)

Le CVE2 retient 15 prédicteurs sur 58 et totalise 1 672 règles réparties sur 100 IVE. L’agrégation retenue est la prédiction médiane des IVE si au moins 50% des IVE se sont exprimés.

Les performances de l’IVE XTRACTIS obtenu après Reverse Engineering du CVE2 et sélectionné sur le RMSE en validation après un découpage 67% / 33% de la base synthétique sont les suivantes, comparées au hasard :

log10 (nb of porosities)	Correlation	MAE	RMSE	Refusal
Descriptive Perf. / Training (2,000is. 67%)	0.997	0.02 (0.53%)	0.02 (0.78%)	0 (0.00%)
Predictive Perf. / Validation (2,000is. 33%)	0.997	0.02 (0.55%)	0.03 (0.86%)	0 (0.00%)
Quality of CVE2 Copy / 153 Original Points	0.898	0.17 (5.65%)	0.24 (7.88%)	0 (0.00%)

Real Performance / External Test	0.824	0.20 (6.45%)	0.25 (7.94%)	0 (0.00%)
Performance against chance to beat / External Test P-Value = 0.001 (among 100,000 random permutations)	<i>0.585</i>	<i>0.28 (8.93%)</i>	<i>0.37 (11.89%)</i>	

Tab. 2. Performances de l'IVE XTRACTIS : $\log_{10}(\text{Nombre de porosités})$

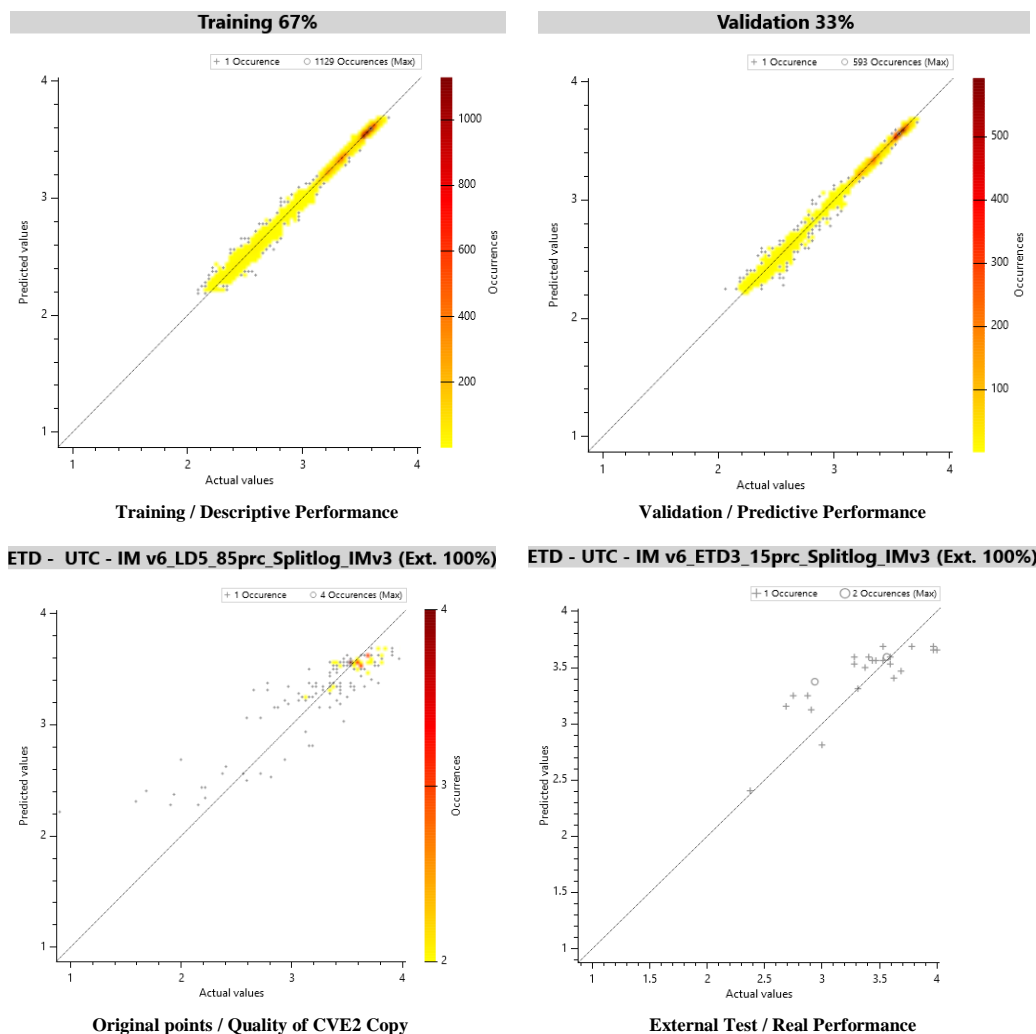


Fig. 2. Graphes de performances de l'IVE2 XTRACTIS : $\log_{10}(\text{Nombre de porosités})$

Sur ETD, ces performances sont plutôt bonnes : 7.94% en RMSE, 0.824 en corrélation. L'intelligibilité du modèle est bonne, avec un score de 3.52/5 : l'IVE n'a gardé que 15 prédicteurs véritablement utiles et il est constitué de 54 règles conjonctives, agrégées en 26 règles disjonctives, avec en moyenne 4.8 prédicteurs par règle. Le score d'intelligibilité d'un modèle prédictif, quelle que soit la technique de génération utilisée, est exposé par exemple dans l'annexe 1 de [8].

Il est ensuite nécessaire de réaliser un benchmark contre une IA non-linéaire très performante. Les performances du **CVE2 Boosted Tree** sélectionné sur le RMSE en validation après un 20×5 folds sont les suivantes :

log₁₀ (nb of porosities)	Correlation	MAE	RMSE
Descriptive Perf. / Training (2,000mls. 20sp × 5f × 80%)	0.985	0.07 (2.20%)	0.12 (3.97%)
Predictive Perf. / Validation (2,000mls. 20sp × 5f × 20%)	0.854	0.19 (6.06%)	0.27 (8.67%)

Real Perf. / External Test	0.834	0.19 (5.97%)	0.23 (7.57%)
----------------------------	-------	--------------	---------------------

Tab. 3. Performances du CVE2 Boosted Tree : $\log_{10}(\text{Nombre de porosités})$

Le CVE2 Boosted Tree retient les 58 prédicteurs, et totalise 159 440 règles réparties sur 100 IVE. L'agrégation retenue est la moyenne des prédictions.

Les performances de l'**IVE Boosted Tree** obtenu en appliquant la stratégie du meilleur CVE2 Boosted Tree à 100% de LD sont les suivantes :

log10 (nb of porosities)	Correlation	MAE	RMSE
Descriptive Perf. / Training (100%)	0.997	0.01 (0.35%)	0.04 (1.23%)
Real Perf. / External Test	0.872	0.16 (5.18%)	0.22 (7.02%)

Tab. 4. Performances de l'IVE2 Boosted Tree : $\log_{10}(\text{Nombre de porosités})$

Le modèle Boosted Tree a également de bonnes performances réelles sur ETD : 7.02% en RMSE, 0.872 en corrélation. Il n'est en revanche pas du tout intelligible (0.00/5) puisqu'il conserve les 58 prédicteurs et est constitué d'une chaîne de 116 arbres totalisant 2 289 règles, soit 15 fois le nombre de points d'apprentissage !

Une des forces majeures d'XTRACTIS réside dans l'intelligibilité des règles générées. Par exemple, la règle suivante illustre comment le modèle prédit le nombre de porosités avec une grande clarté et simplicité (les valeurs sont normalisées par la valeur maximale de chaque variable) :

Expression de la règle R9 :

Si la valeur max de P4 est plus petite qu'environ 73

Et la valeur moyenne de P24 est plus petite qu'environ 540

Et D3 est environ entre 750 000 et 766 000

Et la valeur moyenne de P21 est d'environ 19.71

Alors le nombre de porosités dans la couche du réservoir est faible ($10^{2.36} = 229$).

Sachant que P4 représente la tension d'un toron de fibre carbone, P24 est la vitesse de rotation du mandrin, D3 est le volume de la couche d'enroulement et P21 est l'angle d'ouverture du doctor blade dans le bain de résine. Cette règle démontre la capacité de XTRACTIS à exprimer des relations complexes sous une forme accessible et compréhensible par l'expert(e), facilitant ainsi l'interprétation des mécanismes sous-jacents influençant la qualité du réservoir. Cette transparence est un atout majeur par rapport à des modèles opaques, tels que Boosted Tree, qui, bien que performants, manquent d'une telle clarté et sont difficiles à interpréter, donc à auditer et certifier.

4. Conclusion

Les résultats obtenus avec XTRACTIS illustrent parfaitement l'équilibre entre haute performance et intelligibilité du modèle, malgré une base de données relativement réduite. L'Expert Virtuel Individuel (IVE) XTRACTIS a montré une bonne performance sur le jeu de test externe (ETD), avec une corrélation de 0.824 et un RMSE de 7.94%, ce qui démontre une robuste capacité de généralisation et de prédiction au-delà des données d'entraînement. Ce modèle unitaire présente une bonne intelligibilité (3.52/5) : il est constitué de 54 règles conjonctives, agrégées en 26 règles disjonctives, n'utilisant que 15 prédicteurs parmi 58, avec en moyenne 4.8 prédicteurs par règle.

Sur ce problème de modélisation, l'IA Raisonnante Générale XTRACTIS montre qu'elle est capable d'atteindre des performances similaires au Boosted Tree. Ces performances pourront être améliorées grâce à l'augmentation des cas de référence, à l'extraction d'autres prédicteurs potentiels

des signaux de mesure et à la correction des valeurs 0 dans les signaux bruts qui correspondraient en réalité à des absences de mesures (*missing value*).

L'intelligibilité des modèles XTRACTIS pourra être exploitée pour pratiquer une **induction corrective** : l'analyse des règles peut permettre d'améliorer la qualité de la base en détectant et corrigeant des erreurs dans les points de référence, favorisant ainsi l'induction de futurs modèles plus performants.

Références

- [1] B. Patamaprom, «Conception et durabilité de réservoirs en composites destinés au stockage de l'hydrogène,» Ecole Nationale Supérieure des Mines de Paris, 2014.
- [2] M. Azeem et al, «Application of Filament Winding Technology in Composite Pressure Vessels and Challenges: A Review,» *Journal of Energy Storage*, 2022.
- [3] S. Srebrenkoska et al, «Effect of Process Parameters on Thermal and Mechanical Properties of Filament Wound Polymer-Based Composite Pipes,» 2023.
- [4] B. Meemary, *études expérimentales et analyse numérique de l'endommagement de composites thermoplastiques fabriqués par enroulement filamentaire et intégrant des fibres optiques*, 2024.
- [5] W. Qi et al, « Prediction of void growth and fiber volume fraction based on filament winding process mechanics,» 2020.
- [6] Z. Zalila, *Contribution à une théorie des Relations Floues d'ordre N. Thèse de Doctorat en Contrôle des Systèmes/Mathématiques, D 573, Université de Technologie de Compiègne, Compiègne, France, 474 p, Mars 1993.*
- [7] Z. Zalila, «Essai | Votre Intelligence Artificielle, Connexionniste ou Cognitiviste Floue Augmentée ? INTELLITECH [intelligent technologies], v1.4, France, 26sp, <https://xtractis.ai/white-papers/>,» Septembre 2017.
- [8] Z. Zalila, «Idagrai Labs & Xtractis, XTRACTIS® the General Reasoning AI for Trusted Decisions. Use Case #29 | Naval Security: Identification of Underwater Sounds (Virtual Golden Ear) – Benchmark vs. Logistic Regression, Random Forest, Boosted Tree & Neural Network.,» INTELLITECH [intelligent technologies], v2.0, 6p. <https://xtractis.ai/use-cases/>, France, June 2025.