



MEMOIRE DE FIN D'ETUDES DE MASTER ACADEMIQUE

Filière : **GÉNIE DES PROCÉDÉS**

Option : **GÉNIE DES PROCÉDÉS DE L'ENVIRONNEMENT**

THÈME

**Modélisation de la consommation de l'énergie d'électricité de
l'unité de production de CO₂ de la rayanox par IA**

Présenté par

- 1- BOURROUBEY Afif
- 2- BOURROUBEY Lakhder

Soutenu le 28/06/ 2026 devant le jury composé de :

Président(e) :	OUADJENIA Fatima	MCA	Université de Mostaganem
Examinatrice :	HADDOU Nabila	MCB	Université de Mostaganem
Rapporteur(e) :	BENIDRIS Elbatoul	MCB	Université de Mostaganem

Année Universitaire 2025/2026

DÉDICACE

C'est avec un profond sentiment de fierté et de gratitude que nous dédions ce modeste travail aux êtres qui illuminent notre vie :

À nos très chers parents, paix à leurs âmes

Modèles absolus de sacrifice, d'amour inconditionnel et de bienveillance. Que ce travail soit le reflet de notre immense reconnaissance et le témoignage de notre admiration éternelle pour tout ce que vous nous avez transmis.

À nos frères, sœurs et à nos familles respectives,

Pour votre présence réconfortante et votre patience infinie durant les moments de doute et de rigueur liés à ce projet. Vos encouragements ont été notre plus grand moteur, et nous vous en sommes profondément reconnaissants.

À nos chers amis,

En gage de notre amitié sincère, de notre affection profonde et de notre fidélité indéfectible.

REMERCIEMENTS

On tient à exprimer notre profonde gratitude à l'ensemble de l'équipe de **RAYANOX** pour l'accueil chaleureux et l'accompagnement bienveillant dont on a bénéficié tout au long de notre période de stage.

On s'adresse nos remerciements au directeur général de Rayanox de nous avoir accueilli au sein de son entreprise et de nous avoir permis de vivre une expérience professionnelle riche et constructive.

On remercie particulièrement notre encadrante, Madame Benidris El'batoul, maitre de conférences 'B' à l'université de MOSTAGENM pour son accompagnement, sa disponibilité, ses conseils précieux et le temps qu'elle nous a consacré tout au long de cette période de stage.

On tient à exprimer notre profonde gratitude à madame Dr. OUADJENIA Fatima, maître de conférences 'A' à l'université de Mostaganem d'avoir accepté de présider le membre de jury et d'évaluer notre travail.

On tient également à remercier madame Dr. HADDOU Nabila d'avoir accepté d'examiner ce modeste travail.

Un remerciement particulier pour **l'ensemble des ingénieurs** des unités de production et HSE, qui nous a accompagné sur le terrain, partagé leurs connaissances et consacré leur temps pour nous expliquer les différentes étapes du processus de fabrication.

Enfin, ce stage a été pour nous une étape importante, tant sur le plan professionnel que personnel.

Notre gratitude la plus profonde s'adresse à nos familles respectives pour leur soutien indéfectible, tant sur le plan moral que matériel, ainsi que pour les innombrables sacrifices consentis pour mener à bien nos études.

Liste des figures

Figure I.1. Gaz d'oxygène utilisé pour l'oxygénothérapie et l'enrichissement du flux d'air en milieu médical	20
Figure I.2. Gaz d'azote (N ₂) utilisé pour les opérations de découpe, de purge, de refroidissement et de congélation (Rayanox, n.d.).....	21
Figure I.3. Bouteille d'acétylène (C ₂ H ₂) utilisée dans les procédés industriels (Rayanox, n.d.)	22
Figure I.4. Glace carbonique (CO ₂ solide) utilisée dans les applications industrielles (Rayanox, n.d.).....	23
Figure I.5. L'entrée principal de l'administration (Rayanox, n.d.)	25
Figure I.6. Schéma de procédé de l'unité de séparation d'air cryogénique (Zerrouki, M, s. d.).	27
Figure I.7. Schéma de procédé de production de CO ₂ . (Zerrouki, M, s.d.)	31
Figure I.8. Bouteille de dioxyde de carbone (CO ₂) utilisée dans les applications industrielles (Rayanox, s.d.).	32
Figure I.9 .Caractéristiques de stockage de N ₂ O.....	37
Figure I.10. Vue des réservoirs de stockage des gaz industriels au sein de Rayanox (Photo ...	39
Figure I.11. Opération de remplissage et de conditionnement des bouteilles de gaz industriels dans l'unité de production de Rayanox (Rayanox, n.d.)	40
Figure I.12. Illustration du système QHSE intégrant la Qualité, l'Hygiène, la Sécurité et l'Environnement dans les milieux industriels (Tour de France des Alternatives, s.d.)...	41
Figure II.1. Histoire de l'IA Beyrouiti, G., & Cotton, R. (s. d.).....	46
Figure II.2. Types de Machine Learning. ResearchGate GmbH (2024)	50
Figure II.3. Apprentissage Supervisé. Blent (2026).....	51
Figure II.4. Différents modèles de régression linéaire simple vs multiple	52

Figure II.5. Architecture séquentielle de l'algorithme XGBoost et processus de correction des erreurs résiduelles. (Adapté de GeeksforGeeks, 2025).	55
Figure II.6. Apprentissage non Supervisé (Réalisation personnelle)	56
Figure II.7. Apprentissage semi-supervisé (IONOS ,2026).	57
Figure II.8. Apprentissage par renforcement.	57
Figure III.1. Variables aberrants avant le nettoyage	71
Figure III.2. Dataset après nettoyage des valeurs aberrantes	71
Figure III.3. matrice de corrélation	73
Figure III.4. Distribution des résidus et Q-Q plot des résidus.	74
Figure III.5. Valeurs réelle vs valeurs prédite par algorithme régression vs Random forest....	75
Figure III.6. Valeurs réelle vs valeurs prédite par algorithme régression vs Random forest...	76
Figure III.7. Suivre de la consommation de l'électricité réelle vs modèle (XG BOOST)	78
Figure III.8. Suivre de la consommation de l'électricité réelle vs modèle (Random Forest)...	79

Liste des tableaux

Tableau II.1: Les types d'intelligence artificielle.....	48
Tableau II.2.Les métriques clés.....	60
Tableau II.3 .Tableau des domaines d'application potentiels de l'IA.	62
Tableau III.1.Jeux de données collectées lors de la génération lors de la production de CO ₂ . 69	
Tableau III.2.Métriques des algorithmes.....	77

List des abréviations

Abréviations	Signification
Ar	Argon
ASU	Air Separation Unit
BPF	Bonnes Pratiques de Fabrication
BT	Basse Température
C₂H₂	Acétylène
CaC₂	Carbure de Calcium
CaCl₂	Chlorure de Calcium
Ca(OH)₂	Hydroxyde de Calcium
CO	Monoxyde de Carbone
CO₂	Dioxyde de Carbone
GPL	Gaz de Pétrole Liquéfié
H₂O	Eau
H₂S	Sulfure d'Hydrogène
IA	Intelligence Artificielle
IRM	Imagerie par Résonance Magnétique
ISO	International Organization for Standardization
LAR	Liquid Argon
LIN	Liquid Nitrogen
LOX	Liquid Oxygen
MEA	Monoéthanolamine
ML	Machine Learning
N₂	Azote
N₂O	Protoxyde d'Azote
NH₃	Ammoniac
NH₄NO₃	Nitrate d'Ammonium
NO_x	Oxydes d'Azote
O₂	Oxygène
PH₃	Phosphine
QHSE	Qualité, Hygiène, Sécurité et Environnement
RH	Ressources Humaines
SO₂	Dioxyde de Soufre
RF	Random Forest
XGB	Extreme Gradient Boosting
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
R²	Coefficient de Détermination

Résumé

Face à la forte augmentation de la consommation d'électricité dans de nombreux secteurs, notamment dans le secteur industriel, et à son impact croissant sur la durabilité de la production et le bon fonctionnement de l'activité industrielle, il est devenu nécessaire de trouver des solutions à cette consommation croissante, et ce, grâce à la modélisation de cette consommation d'énergie afin d'améliorer les performances et de développer les processus industriels.

Dans ce contexte, ce travail se concentre sur la modélisation de la consommation d'électricité d'une unité de production de dioxyde de carbone (CO_2) au sein de la société RAYANOX (SARL RAYANOX).

Cette recherche vise à analyser et à modéliser l'influence de certains paramètres de fonctionnement, notamment la pression, la production de dioxyde de carbone et la consommation de GPL, sur la consommation d'électricité de l'unité. Pour atteindre cet objectif, des données industrielles réelles ont été collectées, prétraitées, puis analysées afin de déterminer les relations entre les différentes variables du processus.

Cette recherche s'appuie sur l'utilisation de techniques de science des données et d'apprentissage automatique. L'application de modèles de régression linéaire, de forêt aléatoire et de XGBoost a permis de développer des modèles prédictifs capables d'estimer la consommation d'énergie et de contribuer à l'amélioration du fonctionnement de l'unité.

Mots-clés : *modélisation énergétique, intelligence artificielle, machine Learning, production de dioxyde de carbone.*

Abstract

Given the sharp rise in electricity consumption across many sectors, particularly in the industrial sector, and its growing impact on the sustainability of production and the smooth running of industrial operations, it has become necessary to find solutions to this rising consumption by modelling energy use in order to improve performance and develop industrial processes.

Against this backdrop, this study focuses on modelling the electricity consumption of a carbon dioxide (CO₂) production unit at RAYANOX (SARL RAYANOX).

This research aims to analyse and model the influence of certain operating parameters – notably pressure, carbon dioxide production and LPG consumption —on the unit's electricity consumption. To achieve this objective, real-world industrial data was collected, pre-processed and then analysed to determine the relationships between the various process variables.

This research draws on data science and machine learning techniques.

The application of linear regression, random forest and XGBoost models has enabled the development of predictive models capable of estimating energy consumption and contributing to the improvement of the unit's operation.

Keywords: energy modelling, artificial intelligence, machine learning, carbon dioxide.

في ظل تزايد الكبير في استهلاك الكهرباء في عديد المجالات خاصة ضمن المجال الصناعي، و تأثيره المتزايد الاستدامة الانتاجية و سير الحسن للنشاط الصناعي ، اصبح من ضروري ايجاد حلول لهذا الاستهلاك المتزايد ، و ذلك من خلال نمذجة استهلاك هذه الطاقة لتحسين الأداء وتطوير العمليات الصناعية. في هذا السياق، يركز هذا العمل على نمذجة استهلاك SARL RAYANOX) في شركة ريانوكس (CO_2) الكهرباء لوحدة إنتاج ثاني أكسيد الكربون ..

يهدف هذا البحث إلى تحليل و نمذجة تأثير بعض معايير التشغيل، بما في ذلك الضغط وإنتاج ثاني أكسيد الكربون، واستهلاك غاز البترول المسال، على استهلاك الكهرباء في الوحدة. ولتحقيق هذا الهدف، جُمعت بيانات صناعية واقعية، وخضعت للمعالجة المسبقة، ثم خُللت لتحديد العلاقات بين متغيرات العملية المختلفة.

XGBoost يعتمد هذا البحث على استخدام تقنيات علم البيانات والتعلم الآلي. بتطبيق نماذج الانحدار الخطي، والغابة العشوائية، و لتطوير نماذج تنبؤية قادرة على تقدير استهلاك الطاقة والمساهمة في تحسين تشغيل الوحدة.

الكلمات المفتاحية: نمذجة الطاقة ، الذكاء الاصطناعي، التعلم الآلي، ثاني أكسيد الكربون،

Table des matières

DÉDICACE	ii
REMERCIEMENTS	iii
Liste des figures	iv
Liste des tableaux	vi
List des abréviations	7
Résumé	8
Abstract	9
ملخص :	10
Introduction générale	15
Chapitre I : Présentation	17
de l'entreprise RAYANOX	17
I. Chapitre I : Présentation de l'entreprise RAYANOX	18
I.1 Introduction	18
I.2 Présentation générale de l'entreprise	18
I.2.1 Situation géographique de l'entreprise	19
I.2.2 Importance stratégique et logistique du site.....	19
I.3 Activités et domaines d'intervention	20
I.3.1 Gaz médicaux	20
I.4 Gaz industriels.....	21
I.5 Gaz agroalimentaires	22
I.6 Organisation de l'entreprise	23
I.6.1 Direction Générale	23
I.6.2 Direction Production	23
I.6.3 Direction QHSE.....	24
I.6.4 Direction des Ressources Humaines (RH)	24
I.6.5 Direction Finance	24
I.6.6 Direction Commerciale.....	24
I.6.7 Laboratoire de contrôle de qualité.....	25
I.6.8 Service de maintenance	25
I.7 Les unités de production	26
I.8 Unité de séparation de l'air (ASU)	26
I.9 Fiche technique des produits	27

I.10	Processus de production.....	29
a)	Compression de l'air atmosphérique :	29
b)	Pré refroidissement :	29
c)	Purification de l'air :	29
d)	Refroidissement cryogénique et détente dans la turbine :	29
e)	Distillation fractionnée (colonne de séparation) :	29
f)	Extraction et réchauffage des produits :	30
g)	stockage des effluents liquides :	30
I.11	Unité de production de dioxyde de carbone (CO ₂) :	30
I.11.1	Fiche technique des produits :	31
I.12	Processus de production :	33
1.	Génération des fumées riches en CO ₂ (chaudière) :	33
2.	Rinçage et tour de lavage à l'eau :	33
3.	Absorption du CO ₂ dans la colonne à MEA :	34
4.	Régénération de la MEA et libération du CO ₂ :	34
5.	Pré-refroidissement et compression :	35
6.	Séchage par tamis moléculaire et désulfurisation :	35
7.	Liquéfaction de CO ₂ :	36
8.	Stockage de CO ₂ :	36
I.13	Unité de production de NO ₂ :	36
I.13.1	Fiche technique de produit :	36
I.14	Unité de production d'Acétylène :	38
I.14.1	Fiche technique des produits :	38
I.15	Remplissage des bouteilles :	39
I.15.1	Unité de remplissage et de mélange des gaz.....	39
I.16	Politique Qualité, Hygiène, Sécurité et Environnement (QHSE)	40
	Conclusion	42
	Références Bibliographiques	43
	Chapitre II. L'intelligence artificielle.....	44
	et leurs applications dans le domaine industriel.....	44
II.	Chapitre II. L'intelligence artificielle et leurs applications dans le domaine industriel	45
II.1	Introduction	45
II.2	Définition de l'intelligence artificiel.....	45

II.3	Historique de l'IA	46
II.4	L'objectif de l'intelligence artificielle.....	46
II.5	Les domaines d'application de l'IA	47
II.6	Les types d'intelligence artificielle.....	47
II.6.1	IA faible (Narrow AI)	48
II.6.2	IA forte.....	48
II.6.3	IA générale (AGI) et super-intelligence.....	49
II.7	Les principaux modèles et architectures d'IA	49
II.7.1	Le machine Learning ou apprentissage automatique	49
II.7.2	Le fonctionnement de la machine Learning.....	49
II.8	Les différents types de Machine Learning	50
II.8.1	Apprentissage Supervisé	50
II.8.2	Régression linéaire	51
II.8.3	Régression Linéaire Simple.....	51
II.8.4	Régression Linéaire Multivariable	52
II.8.5	MAE (Mean Absolute Error - Erreur Absolue Moyenne)	53
II.8.6	RMSE(Root Mean Squared Error - Erreur Quadratique Moyenne).....	53
II.8.7	Forêt Aléatoire (Random Forest).....	54
II.8.8	Algorithme XGBoost	54
II.8.9	Apprentissage non Supervisé.....	55
II.8.10	Apprentissage semi-supervisé	56
II.8.11	L'apprentissage par renforcement.....	57
II.8.12	Deep Learning	58
II.9	LLM et modèles multimodaux.....	58
II.10	Comment entraîner et déployer une intelligence artificielle.....	58
II.11	Comment évaluer la performance d'une IA	59
II.12	L'IA enjeu pour l'industrie	60
II.13	Travaux antérieurs de l'application de l'intelligence artificielle dans le domaine industrielle	
	63	
a)	Optimisation des Processus.....	63
b)	Maintenance Prédictive	63
c)	Contrôle de Qualité.....	63
d)	Robotique et Automatisation	64
e)	Supply Chain Management.....	64

Références Bibliographiques	65
CHAPITRE III. Résultats et discussion de la	67
Modélisation de la consommation énergétique	67
de l'unité de la production de CO ₂	67
III. CHAPITRE III. Résultats et discussions de la modélisation de la consommation énergétique de l'unité de la production de CO ₂	68
III.1 Introduction	68
III.2 Collecte des données	68
III.2.1 Nettoyage des valeurs manquantes	70
III.2.2 Nettoyage des valeurs aberrantes.....	70
III.3 Résultats du Nettoyage	72
III.4 Matrice de corrélation des variables (Heat map)	73
III.5 Graphique Q-Q Plot (À droite)	75
Conclusion générale	80
Recommandations :	81
Annexe	82

Introduction générale

Bien que la capacité de réflexion ait longtemps été considérée comme une prérogative exclusivement humaine, les ordinateurs contemporains possèdent désormais la faculté d'apprendre à interpréter leur environnement, de communiquer, de prédire et de créer de manière autonome. Cette évolution majeure de l'intelligence artificielle (IA) est étroitement liée à l'expansion massive des données mondiales. L'intelligence artificielle vise ainsi à résoudre des problèmes complexes associés à diverses activités humaines telles que la prise de décision, le diagnostic, l'analyse des données et l'interprétation du langage (Haton, 2021).

Dans le secteur manufacturier, cette prolifération massive de l'information a conduit à l'émergence du modèle de l'Industrie 4.0 ainsi que de la *Data Science* (science des données), qui s'est imposée ces dernières années comme une discipline charnière combinant les statistiques, l'apprentissage automatique et la visualisation des données. En opérant à l'intersection des technologies de l'information (IT) et des technologies opérationnelles (OT), la *Data Science* transforme les flux de données brutes, collectées en temps réel par les capteurs industriels, en les nettoyant, les analysant et les convertissant en informations prédictives exploitables. Elle applique également des méthodologies avancées fondées sur le *Machine Learning* et le *Deep Learning* afin d'appréhender le comportement des systèmes et des processus de production ainsi que les facteurs qui les influencent. L'ensemble de ces approches vise à optimiser le fonctionnement global, à améliorer les performances de production et à rationaliser la consommation d'énergie.

Dans ce contexte, le complexe SARL RAYANOX constitue un site industriel stratégique fortement numérisé, caractérisé par une haute densité de données. Ce complexe génère quotidiennement un volume important d'informations provenant de multiples sources, relatives à la production, aux paramètres de fonctionnement (pression et débit) ainsi qu'aux modes de consommation énergétique, notamment l'énergie électrique. L'intégration de la *Data Science* dans cette structure, à travers des modèles et des algorithmes d'apprentissage automatique, permet de passer d'une gestion réactive à une gestion proactive, facilitant ainsi la réalisation d'analyses prédictives précises capables d'améliorer l'efficacité énergétique globale des équipements.

L'objectif principal de ce mémoire est de modéliser la consommation électrique de l'unité de production de dioxyde de carbone (CO₂) de l'entreprise SARL RAYANOX, à partir de données

réelles de fonctionnement. Pour ce faire, des techniques d'apprentissage automatique, notamment la Régression Linéaire, *Random Forest* et *XGBoost*, seront implémentées afin d'estimer la consommation énergétique de l'unité étudiée.

Ce travail de fin d'étude est articulé autour de trois chapitres :

- **Le premier chapitre** est dédié à la présentation détaillée de l'entreprise SARL RAYANOX, de ses installations industrielles, de ses différentes unités de production ainsi que de la nature des gaz produits.
- **Le deuxième chapitre** est consacré au cadre théorique de l'intelligence artificielle. Nous y aborderons ses concepts fondamentaux, ses domaines d'application, ainsi que l'apprentissage supervisé et ses algorithmes, le tout complété par un état de l'art des travaux antérieurs dans le domaine industriel.
- **Le troisième chapitre** est dévolu à la partie pratique. Après le traitement et le nettoyage des données industrielles, nous étudierons les corrélations existantes entre les différentes variables. Nous intégrerons ensuite ces valeurs dans les algorithmes de modélisation afin de prédire la consommation électrique à l'aide de la régression linéaire multivariable, de *Random Forest* et de *XGBoost*.
- **Et** en clôturant ce manuscrit par une conclusion et les principales perspectives issues de ce travail

En résumé, ce travail de recherche s'inscrit au cœur de la transition numérique et énergétique de l'industrie moderne. En appliquant les outils de la *Data Science* aux paramètres opérationnels réels du site de la société SARL RAYANOX, cette étude vise à démontrer que la convergence entre l'ingénierie des procédés et l'intelligence artificielle constitue un levier stratégique pour prévoir avec précision la consommation d'électricité, rationaliser son utilisation et soutenir la prise de décision en faveur d'une production plus durable.

Chapitre I : Présentation de l'entreprise RAYANOX



I. Chapitre I : Présentation de l'entreprise RAYANOX

I.1 Introduction

Ce premier chapitre se dédie à l'exploration de l'environnement industriel de notre recherche. Nous offrons un aperçu détaillé de la société à responsabilité limitée RAYANOX, en soulignant son historique, son positionnement stratégique dans la wilaya d'Oran et sa structure organisationnelle. En outre, nous examinerons les diverses unités de production au sein du complexe, leurs capacités d'opération et la politique QHSE d'ensemble qui supervise le fonctionnement de l'entreprise.

I.2 Présentation générale de l'entreprise

L'entreprise algérienne SARL RAYANOX se spécialise dans la fabrication, l'emballage et la distribution de gaz à usage médical, industriel et agroalimentaire. Créée en 2011 par le Dr Amine Tazi, un chirurgien cardiologue diplômé de la faculté de médecine de Montpellier, et dirigée conjointement avec Mme Amina Tazi, pharmacienne, la société a su se positionner progressivement comme un intervenant majeur dans le secteur des gaz en Algérie.

Effectivement, c'est en 2021 que l'entreprise a lancé ses opérations industrielles. Il est impératif de souligner que ce lancement a coïncidé avec une période sanitaire extrêmement critique à l'échelle mondiale et nationale, marquée par la pandémie de la COVID-19. Durant cette crise majeure, caractérisée par une pénurie aiguë d'oxygène médical indispensable à la survie des patients, l'ouverture des portes de RAYANOX a constitué un tournant stratégique. En effet, face à la crise sanitaire, le ministère de l'Industrie a annoncé le lancement d'une « action sectorielle d'urgence en vue de la réquisition de tous les moyens disponibles » pour augmenter la production d'oxygène médical, insistant sur la nécessité d'assurer un approvisionnement « régulier et continu » des structures de santé, y compris les week-ends (Zou, K. H., et al, Algérie, 2021). Grâce à ses installations de séparation d'air de grande capacité, l'entreprise a pleinement intégré cet effort national en répondant en urgence à la demande massive des établissements hospitaliers algériens, contribuant ainsi directement à la gestion de cette crise de santé publique.

RAYANOX possède un complexe industriel de pointe qui s'étend sur plus de 24 000 m², et une surface bâtie de 10 000 m², qui abrite diverses unités de production spécialisées dans la séparation d'air, la production de protoxyde d'azote, de dioxyde de carbone et d'acétylène, ainsi

que le remplissage et le mélange des gaz médicaux et industriels. Ses produits ciblent principalement les domaines de la santé, de l'industrie agroalimentaire, de l'énergie, de l'pétrochimie, de l'environnement et de l'agriculture (**Rayanox, n.d.**).

Cette installation aura aussi la capacité de produire des gaz médicaux tels que : le CO₂ avec un rendement quotidien de 5 000 kg, le N₂O avec une production quotidienne de 1 500 kg et l'acétylène également à 1 500 kg par jour (**Algerie Eco, 2021**) .

L'entreprise possède des installations industrielles à la pointe de la technologie qui facilitent la production de divers gaz destinés aux établissements médicaux, à l'industrie et au secteur agroalimentaire. Avec son expertise, ses équipements de pointe et son système intégré de gestion QHSE, RAYANOX se distingue comme un acteur majeur dans son domaine en Algérie.

Rayanox met un accent particulier sur la qualité, l'hygiène, la sécurité et la protection de l'environnement via un système de gestion intégré QHSE. Elle a ainsi pu décrocher plusieurs certifications mondiales, y compris ISO 9001, ISO 22000 et ISO 13485, qui soulignent son dévouement pour l'amélioration constante et le contentement de ses clients.

1.2.1 Situation géographique de l'entreprise

L'entreprise RAYANOX se trouve dans la zone industrielle de Chehairia, Lot 353, Lot 07, à Bethioua, au sein de la wilaya d'Oran. Cette position stratégique, près du site pétrochimique d'Arzew et des installations portuaires, procure à la société des conditions idéales pour se procurer des matières premières et pour déployer ses produits sur les marchés national et mondial.

1.2.2 Importance stratégique et logistique du site

L'emplacement de la zone industrielle de Bethioua n'est pas le fruit du hasard ; il s'aligne avec une stratégie de grappes industrielles (Industrial Cluster). La position stratégique de RAYANOX à proximité du complexe pétrochimique d'Arzew lui confère un avantage au sein d'un dense écosystème industriel, favorisant l'établissement de synergies techniques, l'accès facilité aux infrastructures énergétiques et un approvisionnement en matières premières optimisé (**BOUZIDI, 2018**) et il facilite le processus de transport et de distribution.

En outre, du point de vue logistique, les deux façades portuaires (les ports d'Oran et d'Arzew) et le raccordement direct aux principaux axes autoroutiers, y compris l'Autoroute Est-Ouest, représentent des avantages concurrentiels considérables (**BOUZIDI, 2018**) et importantes . Cette infrastructure de transport routier et maritime de haute qualité diminue considérablement les coûts

et les temps de transport des gaz liquéfiés ou comprimés sous haute pression. D'après les fiches logistiques de la société, cela délivre un maximum de sécurité pendant les livraisons et offre une capacité de réponse idéale pour l'approvisionnement des hôpitaux et des complexes industriels dispersés sur l'ensemble du pays, tout en ouvrant des avenues réelles pour l'exportation.

I.3 Activités et domaines d'intervention

RAYANOX se spécialise dans la fabrication, l'emballage et la distribution de gaz médicaux, industriels et agroalimentaires. La société s'implique aussi dans divers services concernant les installations industrielles et les analyses de gaz.

I.3.1 Gaz médicaux

L'oxygène médical (O_2) : Emploi pour l'oxygénothérapie, les manœuvres de réanimation et l'augmentation de la teneur en oxygène de l'air dans les environnements médicaux. Elle doit respecter strictement une pureté égale ou supérieure à 99,5%, conformément à ce qui est prescrit dans la Pharmacopée Européenne (**Ph. Eur., 2020**).



Figure I.1. Gaz d'oxygène utilisé pour l'oxygénothérapie et l'enrichissement du flux d'air en milieu médical

a. Le protoxyde d'azote médical (N_2O) :

Essentiellement utilisé comme agent d'anesthésie et de soulagement de la douleur dans les établissements hospitaliers.



Figure I.2. Gaz d'azote (N_2) utilisé pour les opérations de découpe, de purge, de refroidissement et de congélation (Rayanox, n.d.)

b. Le mélange antalgique TITINOX :

Un produit vedette de la société qui est constitué à parts égales de 50% d'oxygène et de 50% de protoxyde d'azote. Ce mélange gazeux stable est utilisé dans le cadre de l'analgésie d'urgence, les soins pour enfants et lors des opérations mineures ou éprouvantes.

Les mélanges gazeux médicaux spécifiques : Conçus pour les laboratoires d'analyses médicales et les diagnostics fonctionnels.

On utilise ces gaz dans :

- Les soins en milieu hospitalier
- Les opérations chirurgicales
- Les services d'urgence
- Les soins destinés aux enfants
- L'anesthésie et l'analgésie.

I.4 Gaz industriels

RAYANOX fabrique également une variété de gaz industriels employés dans divers domaines industriels, y compris :

- L'oxygène industriel ;
- L'azote ;
- Le dioxyde de carbone (CO_2)
- L'acétylène.

On utilise ces gaz dans les domaines suivants :

- La soudure et la découpe industrielles
- Les méthodes de production
- Les processus de refroidissement
- Les procédures industrielles spécialisées.



Figure I.3. Bouteille d'acétylène (C_2H_2) utilisée dans les procédés industriels (Rayanox, n.d.)

I.5 Gaz agroalimentaires

L'entreprise se charge également de la production et de la distribution de gaz à destination du secteur agroalimentaire. On utilise ces gaz pour :

- La préservation des aliments
- L'emballage en atmosphère régulée
- Le secteur des boissons pétillantes
- Les usages du refroidissement industriel.



Figure I.4. Glace carbonique (CO_2 solide) utilisée dans les applications industrielles (Rayanox, n.d.)

I.6 Organisation de l'entreprise

La structure de Rayanox est composée de plusieurs départements responsables de la gestion administrative, technique et opérationnelle de la société.

I.6.1 Direction Générale

La direction générale est responsable de :

- La gestion d'ensemble de l'entreprise
- Les décisions stratégiques
- L'expansion industrielle et commerciale
- Le monitoring de la performance de la société.

Elle supervise toutes les directions pour garantir le fonctionnement optimal de l'entreprise.

I.6.2 Direction Production

La direction de production est responsable :

- De la gestion des unités de production
- De la surveillance des processus industriels
- De l'utilisation des équipements ;
- Du respect des objectifs de production.

Elle prend soin de garantir la continuité opérationnelle des équipements industriels.

I.6.3 Direction QHSE

Le service QHSE (Qualité, Hygiène, Sécurité et Environnement) est responsable de :

- La vérification de la qualité des produits
- La prévention des dangers industriels
- La mise en œuvre des standards de sécurité
- La sauvegarde de l'environnement
- La gestion des certifications ISO.

I.6.4 Direction des Ressources Humaines (RH)

La direction des Ressources Humaines est responsable de :

- L'embauche de nouveau personnel
- La formation professionnelle des salariés
- La gestion administrative des employés
- Le suivi des qualifications et des autorisations.

Elle s'implique aussi dans l'organisation des formations QHSE.

I.6.5 Direction Finance

La gestion financière de l'entreprise, le contrôle des dépenses et des investissements, la comptabilité et l'examen des coûts de production sont assurés par la direction financière.

Elle participe à la surveillance économique des opérations industrielles.

I.6.6 Direction Commerciale

La gestion commerciale est en charge de :

- La mise en vente des produits
- La gestion des relations client
- Le suivi des transactions
- L'expansion du marché.

Elle est également en charge du suivi des exigences dans les domaines médicaux et industriels.



Figure I.5.L'entrée principale de l'administration (Rayanox, n.d.)

I.6.7 Laboratoire de contrôle de qualité

Le laboratoire d'assurance qualité occupe une position cruciale et stratégique au cœur du complexe Rayanox. Il a la responsabilité d'exécuter les analyses physicochimiques strictes et essentielles pour assurer la conformité, la sécurité et l'homologation de chaque lot de gaz avant sa distribution commerciale. L'accent des analyses est principalement mis sur :

Le niveau de pureté élevé : Contrôle rigoureux de la concentration des gaz (particulièrement l'oxygène médical qui doit être au moins égal à 99,5%) par le biais de méthodes analytiques sophistiquées telles que la chromatographie gazeuse (MESSAOUDI, 2019).

La détection des traces d'humidité et d'impuretés : control of carbon monoxide (CO) and carbon dioxide (CO₂), as well as residual water vapour, which could lower the product's quality or damage storage bottles in accordance with industrial gas analysis specifications (MESSAOUDI, 2019).

Le respect des réglementations : Contrôle des protocoles conformément aux standards pharmaceutiques nécessaires pour les usages thérapeutiques (Ph. Eur., 2020).

I.6.8 Service de maintenance

- La maintenance préventive des équipements
- Le contrôle des compresseurs, réservoirs, conduites et dispositifs de sûreté

- Les actions correctives en cas de défaillance.

I.7 Les unités de production

D'après les données techniques accessibles, Rayanox possède diverses installations industrielles dédiées à la fabrication et à l'emballage des gaz médicaux et industriels.

L'entreprise possède :

- Une unité de séparation de l'air (ASU)
- Une unité de production de protoxyde d'azote
- Une unité de production de CO₂
- Une unité de production d'acétylène
- Une unité de remplissage et de mélange des gaz médicaux et industriels.

On estime que les capacités de production à l'échelle globale sont les suivantes (**Rayanox, n.d**) :

- 100 000 litres d'oxygène par jour
- 2 tonnes de protoxyde d'azote par jour
- 10 tonnes de CO₂ par jour
- 1440 kg d'acétylène par jour
- 5000 bouteilles remplies par jour.

I.8 Unité de séparation de l'air (ASU)

L'entreprise dispose de l'Unité de Séparation de l'Air (ASU) comme l'une de ses installations de production primordiales. Elle garantit la division cryogénique des gaz atmosphériques dans le but de générer principalement (**Zerrouki, M, s. d.**):

- L'oxygène (O₂)
- L'azote (N₂)
- L'argon (Ar).

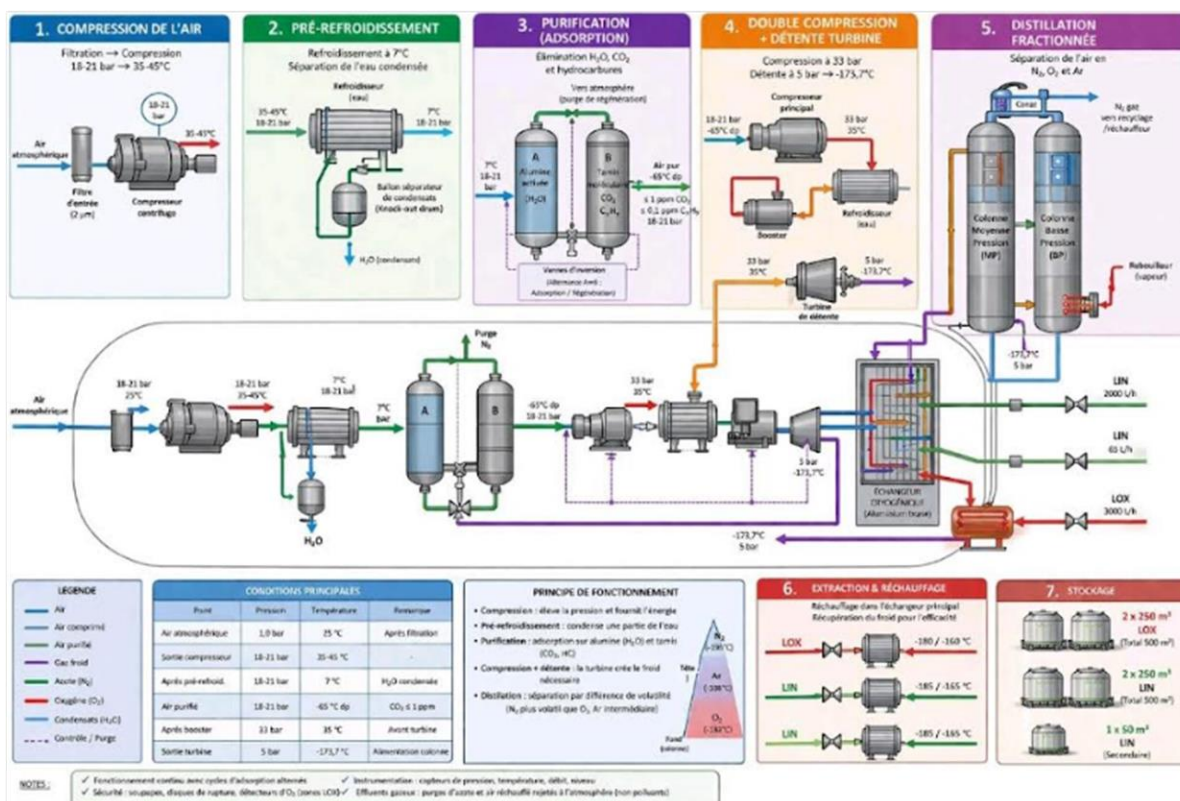


Figure I.6. Schéma de procédé de l'unité de séparation d'air cryogénique (Zerrouki, M, s. d.).

I.9 Fiche technique des produits

➤ Oxygène liquide (LOX)

État : Liquide cryogénique

Température : Environ -183 °C (à pression atmosphérique)

Couleur/odeur : Incolore, inodore

Densité : Environ 1 140 kg/m³ (à l'état liquide)

Pureté typique : 99,5 % à 99,999 % selon application

Stockage : Réservoirs cryogéniques isothermes (sous vide ou perlite)

Pression de stockage : Généralement 1 à 10 bar selon la conception

Utilisations : Combustion, oxycoupage, traitement des eaux, applications médicales (**Zerrouki, M. s. d.**).

➤ Azote liquide (LIN)

État : Liquide cryogénique

Température : Environ -196°C (à pression atmosphérique)

Couleur / odeur : Incolore, inodore

Densité : Environ 807 kg/m^3 (à l'état liquide)

Pureté typique : 99,9 % à 99,999 %

Stockage : Réservoirs cryogéniques isothermes

Pression de stockage : Généralement 1 à 10 bar

Utilisations : Surgélation, inertage, refroidissement, laboratoires, électronique (**Zerrouki, M, s. d.**).

➤ **Argon liquide (LAR)**

État : Liquide cryogénique

Température : Environ -186°C (à pression atmosphérique)

Densité : Environ $1\,394\text{ kg/m}^3$ (à l'état liquide)

Pureté typique : 99,99 % à 99,999 %

Stockage : Réservoirs cryogéniques isothermes

Utilisations : Soudage (protection de l'arc), métallurgie, électronique, éclairage, isolation thermique (double vitrage) (**Zerrouki, M, s. d.**).

➤ **Caractéristiques générales de stockage**

Isolation : Double paroi sous vide ou perlite

Évaporation naturelle : 0,3 à 1 % par jour selon capacité

LOX : risque d'incendie/explosion (comburant)

LIN et LAR : risque d'asphyxie (déplacement d'oxygène) – l'argon est plus dense que l'air, peut stagner en bas des locaux

Remplissage/extraction : Lignes cryogéniques isolées, vannes spécifiques BT

I.10 Processus de production

a) Compression de l'air atmosphérique :

Le compresseur principal comprime l'air atmosphérique préfiltré jusqu'à atteindre une pression de 18 à 21 bars, tout en maintenant une température de sortie entre 35 et 45 °C grâce à un système de refroidissement intermédiaire. L'air comprimé est par la suite dirigé vers la phase de pré-refroidissement.

b) Pré refroidissement :

L'air comprimé à une pression de 18–21 bars et à une température de 35–45 °C est rafraîchi jusqu'à 7 °C par le biais d'un système de refroidissement à eau, sans modification notable de la pression. Une portion de la vapeur d'eau se transforme en condensat et est ensuite isolée dans un récipient séparateur avant que l'air soit dirigé vers l'unité de purification.

c) Purification de l'air :

L'air pré-réfrigéré à 7 °C et soumis à une pression de 18-21 bars passe à travers un dispositif d'adsorption comportant deux adsorbants qui fonctionnent en alternance. L'alumine activée supprime l'humidité résiduelle, alors que le CO₂ et une portion des hydrocarbures sont capturés par un tamis moléculaire. Au terme du processus, le niveau de CO₂ est situé en dessous de 1 ppm et le point de rosée reste au-dessous de –65 °C, avec une pression quasiment stable.

d) Refroidissement cryogénique et détente dans la turbine :

L'air épuré est comprimé par un compresseur de surpression jusqu'à atteindre environ 33 bars, à une température approximative de 35 °C. Après avoir été refroidi, il est libéré dans une turbine jusqu'à atteindre 5 bars, ce qui réduit sa température à –173,7 °C. Une portion de l'air se transforme en liquide et fournit par la suite la colonne de distillation cryogénique.

e) Distillation fractionnée (colonne de séparation) :

L'air, légèrement liquéfié à –173,7 °C et 5 bars, est dirigé vers la colonne de distillation où l'azote se condense dans le haut de la colonne et l'oxygène à sa base. Cette dissociation conduit à la production d'environ 3000 litres par heure de LOX, 2000 litres par heure de LIN et un deuxième flux de 65 litres par heure de LIN.

f) Extraction et réchauffage des produits :

Les courants liquides cryogéniques passent par des échangeurs thermiques pour capter le froid et faciliter leur passage vers l'espace de stockage. L'azote liquide est conservé entre -180 °C et -160 °C, pendant que les flux d'azote liquide obéissent au même mécanisme sans une vaporisation totale.

g) stockage des effluents liquides :

Deux réservoirs de 250 m³ chacun serve à stocker l'oxygène liquide, dont la production atteint 3000 L/h. Deux réservoirs de 250 m³ chacun contient l'azote liquide à 2000 L/h, tandis qu'un réservoir spécifique de 50 m³ est prévu pour le flux de 65 L/h. Les réservoirs possèdent des mécanismes d'isolation cryogénique et des systèmes de sûreté garantissant une conservation sûre des produits en phase liquide.

1.11 Unité de production de dioxyde de carbone (CO₂) :

Cette unité se consacre à la collecte, purification, déshydratation et liquéfaction du dioxyde de carbone (CO₂) dérivé des émissions de combustion industrielle.

Dans cette unité, l'absorption chimique à l'aide d'alcanolamines, comme la monoéthanolamine (MEA), est couramment utilisée dans la capture industrielle du dioxyde de carbone à partir des gaz de combustion [Kohl et Nielsen, 2017]. Ce principe technologique constitue la base du fonctionnement de l'unité de Rayanox pour séparer et purifier le CO₂. Le procédé se décompose ensuite en plusieurs opérations unitaires successives.

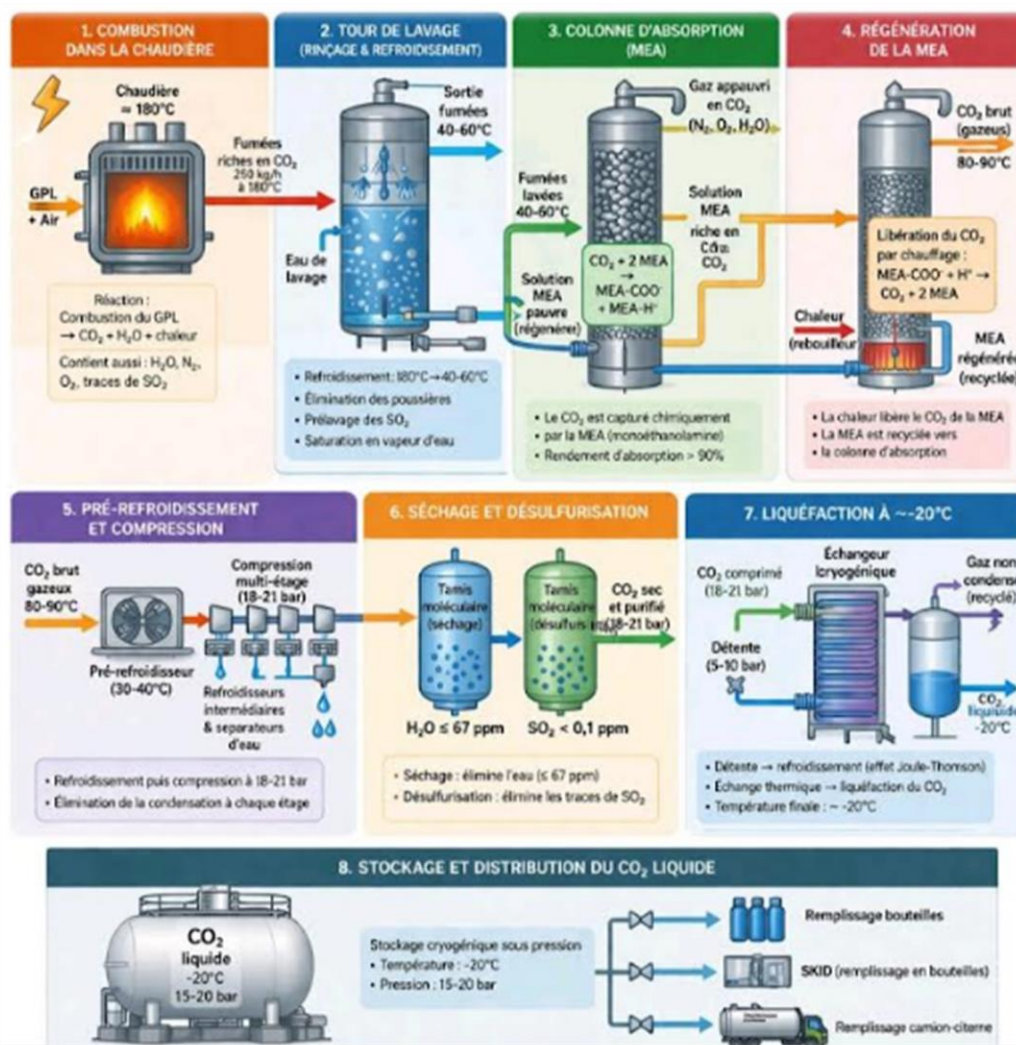


Figure I.7. Schéma de procédé de production de CO_2 . (Zerrouki, M, s.d.)

I.11.1 Fiche technique des produits :

Dioxyde de carbone (CO_2)

- État : Liquide cryogénique (après liquéfaction) ou gazeux sous pression
- Température de liquéfaction : Environ -20°C (à pression élevée)
- Pression de stockage liquide : Typiquement 15 à 20 bar (selon la conception)
- Pureté typique : $\geq 99,9\%$ (après absorption MEA + tamis moléculaire), avec traces d'impuretés (H_2O , SO_2 , hydrocarbures) réduites
- Débit de production : 250 kg/h de CO_2 (indiqué sur le schéma)

- Stockage : Réservoir de stockage liquide (cryogénique ou sous pression)
- Remplissage : En bouteilles (skid bouteilles) ou en camion-citerne
- Utilisations : Industrie agroalimentaire, soudage, traitement d'eau, neutralisation, etc.



Figure I.8.Bouteille de dioxyde de carbone (CO₂) utilisée dans les applications industrielles (Rayanox, s.d.).

➤ **Caractéristiques générales de stockage**

- État du produit : CO₂ liquide (cryogénique sous pression)
- Température de stockage : Environ -20 °C (correspondant à la sortie du liquéfacteur)
- Pression de stockage : Typiquement 15 à 20 bar (pour maintenir le CO₂ à l'état liquide à -20 °C)

➤ **Type de réservoir :**

Réservoir sous pression isotherme (double paroi)

Isolation : vide ou perlite pour limiter les apports de chaleur

- Capacité : Non spécifiée sur le schéma, mais dimensionnée en fonction du débit de production (250 kg/h) et de l'autonomie souhaitée (ex. 24h → 6 000 kg, soit environ 6 m³ de liquide)
- Évaporation naturelle : Faible (< 0,5 % par jour) grâce à l'isolation ; les gaz d'évaporation sont généralement récupérés ou recondensés
- Sécurité : Soupapes de surpression (limite à 25-30 bar)
- Détecteurs de CO₂ (risque d'asphyxie)
- Ventilation des aires de stockage

I.12 Processus de production :

1. Génération des fumées riches en CO₂ (chaudière) :

L'initialisation du processus se fait par la combustion maîtrisée d'un mélange d'air et de gaz de pétrole liquéfié (GPL) dans une chaudière de grande capacité. La réaction de combustion du GPL, qui se compose surtout de propane et de butane, génère de la chaleur, de la vapeur d'eau et du dioxyde de carbone. L'excès d'air est régulé pour maximiser le niveau de CO₂ dans les gaz d'échappement tout en restreignant la production de monoxyde de carbone et d'oxydes d'azote. Les émissions de la chaudière sont libérées à une température approximative de 180 °C, avec un taux de libération de CO₂ de 250 kg/h. On peut exploiter la chaleur émise, par exemple, pour réchauffer des fluides ou générer de la vapeur sous basse pression. Les fumées renferment aussi de l'azote résiduel, de l'oxygène non utilisé, de la vapeur d'eau, et des traces de composés sulfurés (SO₂) dérivant des impuretés du GPL.

2. Rinçage et tour de lavage à l'eau :

Les fumées de haute température (180 °C) sont dirigées vers une colonne de lavage (scrubber) en opposition au flux de l'eau. Cette tour comporte de multiples fonctions :

- Refroidissement rapide : La température des gaz chute de 180 °C à une plage allant de 40 °C à 60 °C. Ce refroidissement préserve l'intégrité de la solution d'amine (MEA) qui serait affectée si la température dépasse 60 °C.
- Élimination des poussières : L'eau capture les particules solides générées par la combustion.
- Saturation en vapeur d'eau : L'air se sature en vapeur d'eau, ce qui est essentiel pour l'efficacité de l'absorption au MEA (la présence de l'eau favorise la réaction chimique).
- Élimination préliminaire du SO₂ : Une portion du dioxyde de soufre, hautement soluble, est

absorbée par l'eau de lavage, diminuant ainsi la quantité requise pour l'étape suivante de désulfurisation.

Après décantation et refroidissement, l'eau de lavage est recyclée et une purge régulière est mise en place pour prévenir l'accumulation de contaminants.

3. Absorption du CO₂ dans la colonne à MEA :

Les gaz sont refroidis (40-60 °C) entrent par la base d'une colonne d'absorption remplie (ou à niveaux), où ils traversent en sens inverse d'une solution aqueuse de monoéthanolamine (MEA) généralement concentrée entre 20 et 30% en poids. La MEA est une amine primaire qui interagit chimiquement avec le CO₂ suivant la réaction :



Cette réaction est réversible et entraîne une libération de chaleur. Cette technologie peut sélectionner vivement capturer le CO₂, aussi en présence d'azote et d'oxygène. Le gaz résiduel est émis à la sortie de la colonne.

Le CO₂ étant appauvri et se composant principalement d'azote, d'oxygène et de vapeur d'eau, il est ensuite rejeté dans l'atmosphère. On récupère la solution de MEA chargée (haute en CO₂) en base de la colonne, avant de la chauffer et de l'acheminer vers la colonne de régénération.

Pour le CO₂, le rendement d'absorption dépasse généralement 90%.

4. Régénération de la MEA et libération du CO₂ :

La solution de MEA chargée, souvent dénommée « riche », est dirigée vers un échangeur de chaleur (économiseur) après quoi elle est injectée dans la partie supérieure d'une colonne de régénération (désorbeur). Un rebouilleur (généralement à la vapeur ou à l'aide des gaz de combustion chauds) sert à réchauffer le fond de cette colonne. La réaction chimique est inversée par la chaleur, libérant le CO₂ sous une forme gazeuse. La MEA régénérée est éjectée à l'arrière et après refroidissement, est transférée vers l'absorbeur. Le CO₂ émis en haut de la colonne est chargé de vapeur d'eau et renferme de petites quantités d'impuretés (comme le SO₂, NH₃, hydrocarbures légers). Sa température atteint approximativement 80-90 °C, et il est orienté vers les processus de compression et de purification. La régénération est ajustée de manière à réduire la détérioration thermique de la MEA (la température du rebouilleur est généralement fixée à 120 °C).

5. Pré-refroidissement et compression :

Le CO₂ brut (qui est gazeux, chaud et humide) est tout d'abord dirigé vers un échangeur de chaleur à air ou à eau pour réduire sa température à environ 30–40 °C. Ce pré-refroidissement condense une portion de l'eau, qui est ensuite retirée dans un séparateur. Le gaz est ensuite comprimé par un compresseur à plusieurs niveaux (généralement 2 ou 3 niveaux), jusqu'à atteindre une pression se situant entre 18 et 21 bars. Entre chaque niveau, un dispositif de refroidissement intermédiaire et un séparateur d'eau se chargent d'éliminer les condensats. La compression du gaz entraîne une hausse de sa température ; les refroidisseurs sont généralement utilisés pour garder la température d'entrée de l'étape suivante en dessous de 40 °C, On opte pour une pression finale de 18-21 bars pour faciliter la liquéfaction du CO₂ à une température raisonnable (environ -20 °C) sans avoir recours à des pressions excessivement hautes, ce qui permet de diminuer la consommation énergétique.

6. Séchage par tamis moléculaire et désulfurisation :

Le CO₂, sous forme comprimée à une pression de 18 à 21 bars et à température ambiante (suite à un refroidissement final), passe à travers un dispositif d'adsorption en deux étages, soit en série, soit en alternance. Le premier lit, qui est composé d'alumine activée ou de tamis moléculaire de type 3A, emmagasine de façon sélective l'humidité résiduelle. Ainsi, la quantité d'eau est diminuée à un niveau ne dépassant pas 67 ppm, limite essentielle pour prévenir la création de glace ou d'hydrates lors de sa liquéfaction future. Le second lit, également appelé lit combiné, est spécifiquement destiné à la désulfurisation : il intègre un tamis moléculaire imprégné d'un agent basique (tels que du charbon actif modifié ou une zéolithe échangeuse d'ions) qui effectue une adsorption chimique des résidus de dioxyde de soufre (SO₂) subsistant après le processus d'absorption MEA. Ce processus en deux étapes (séchage et désulfurisation) assure un CO₂ final d'une pureté très élevée, adapté à des applications délicates (secteur alimentaire, médical). Les deux adsorbeurs opèrent en rotation : l'un est en opération tandis que l'autre subit une régénération via un chauffage à 150-200 °C sous l'action d'un gaz sec (azote ou CO₂ pur). Au terme de ce processus, le CO₂ est intégralement débarrassé de l'eau (≤ 67 ppm) et du soufre (habituellement $< 0,1$ ppm), avant d'être acheminé vers la phase de liquéfaction.

7. Liquéfaction de CO₂ :

Le CO₂ purifié, sec et désulfuré (pression 18–21 bars, température ambiante), est divisé en deux flux. Le flux principal (environ 70 à 80 % du débit total) est dirigé vers un échangeur de chaleur cryogénique. L'autre flux (20 à 30 %) est d'abord détendu à travers une vanne de détente (ou une turbine de détente) pour abaisser sa pression jusqu'à environ 5–10 bars. Cette détente brutale (effet Joule-Thomson) provoque une chute de température du CO₂ gazeux, qui descend jusqu'à environ –50 °C à –60 °C. Ce gaz froid traverse ensuite l'échangeur cryogénique à contre-courant du flux principal, lui cédant ses calories. Le flux principal est ainsi refroidi jusqu'à une température d'environ –20°C . À cette température et sous une pression maintenue à 18–21 bars, le CO₂ se condense partiellement ou totalement en phase liquide. Le mélange de liquide et de vapeur résiduelle entre dans un ballon séparateur (séparateur liquide/vapeur). Le CO₂ liquide (à –20 °C) est soutiré en fond et envoyé vers le stockage. Le CO₂ gazeux non condensé (essentiellement du gaz froid issu de la détente, réchauffé dans l'échangeur) est recyclé en amont du compresseur pour être à nouveau comprimé et mélangé à l'alimentation. Ce procédé auto-réfrigérant évite le recours à un fluide frigorigène externe, améliorant l'efficacité énergétique de l'unité. La température de –20 °C est ainsi atteinte et maintenue de manière stable, avec une pression de liquéfaction adaptée.

8. Stockage de CO₂ :

Liquide Le CO₂ liquide purifié est dirigé vers un réservoir de stockage cryogénique sous pression (double paroi isolée). Ce réservoir maintient le produit à une température d'environ –20 °C et une pression de 15 à 20 bars.

I.13 Unité de production de NO₂ :

I.13.1 Fiche technique de produit :

- Protoxyde d'azote (N₂O)
 - o État : Liquide cryogénique (sous pression) ou gazeux médical/industriel
 - o Température d'ébullition : –88,5 °C (à pression atmosphérique)
 - o Pression de stockage liquide : Typiquement 18 à 24 bars (à température ambiante, autorégulé)
 - o Masse volumique du liquide : Environ 1 230 kg/m³ (à –20 °C, 20 bars)

o Pureté typique : $\geq 99,5\%$ (usage médical ou alimentaire) ; les impuretés contrôlées : NO_x , NH_3 , H_2O

Débit de production : 30 à 60 kg/h (indiqué sur le schéma)

- Matière première : Nitrate d'ammonium (NH_4NO_3) décomposé thermiquement
- Stockage : Réservoir sous pression (Storage tank) après liquéfaction
- Distribution : Remplissage gravimétrique de bouteilles individuelles
- Utilisations : Anesthésique (gaz hilarant), propulseur pour crème chantilly, oxydant en moteur, industrie électronique.

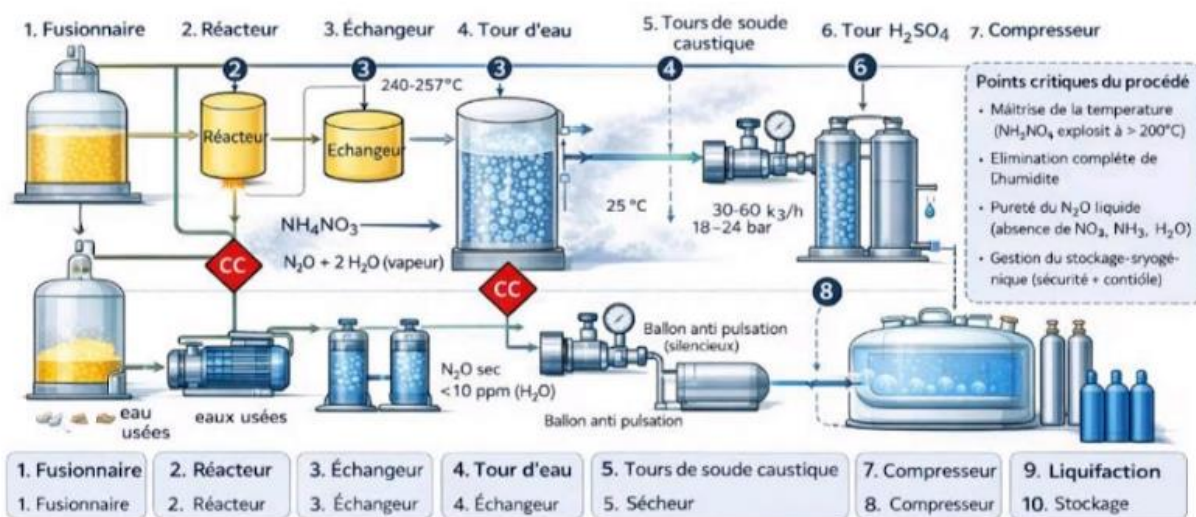


Figure I.9 .Caractéristiques de stockage de N_2O

- État du produit : N_2O liquide (cryogénique sous pression)
 - o Température de stockage : Environ -10°C à -20°C (autogénérée par l'équilibre liquide-vapeur à la pression de stockage)
 - o Pression de stockage : Typiquement 18 à 24 bar (pour maintenir le N_2O à l'état liquide)
 - o Type de réservoir :
 - o Réservoir sous pression isotherme (double paroi)
 - o Isolation : vide ou perlite pour limiter les apports de chaleur

- o Capacité : Non spécifiée sur le schéma, mais dimensionnée en fonction du débit de production (30 à 60 kg/h) et de l'autonomie souhaitée (ex. 24 h → 720 à 1 440kg, soit
- Environ 0,6 à 1,2 m³ de liquide ; réservoir typique de 2 à 10 m³)
 - o Évaporation naturelle : Faible (< 0,5 % par jour) grâce à l'isolation ; les gaz d'évaporation sont généralement récupérés ou recondensés
 - o Sécurité :
 - o Soupapes de surpression (limite à 30 bars)
 - o Détecteurs de N₂O (risque d'asphyxie et effet narcotique)
 - o Ventilation des aires de stockage

I.14 Unité de production d'Acétylène :

Selon **Zerrouki, M. (s.d.)**, cette division se consacre à la fabrication d'acétylène (C₂H₂), un gaz industriel qui est largement utilisé pour les activités de soudage et d'oxycoupage en raison de sa grande capacité calorifique.

I.14.1 Fiche technique des produits :

Acétylène (C₂H₂)

- o État : Gaz comprimé (dissous dans l'acétone pour le stockage en bouteilles) ou gazeux à basse pression
- o Température d'ébullition : -84 °C (à pression atmosphérique)
- o Pression de stockage en bouteille :
Typiquement 17 à 20 bar (à 15 °C) – ici compression à 25 bars puis refroidissement, remplissage à pression adaptée (ex. 20 bar)
- o Masse volumique gazeuse : 1,17 kg/m³ (dans les conditions normales)
- o Pureté typique : ≥ 98 % (acétylène brut après purification), ≥ 99,5 % pour usage technique
- o Matière première : Carbure de calcium (CaC₂) et eau
- o Débit de production : Non spécifié, variable selon la taille du générateur
- o Sous-produit : Résidu de chaux (hydroxyde de calcium, Ca(OH)₂)
- o Utilisations : Soudage oxyacétylénique, synthèse chimique (chlorure de vinyle, caoutchouc), éclairage

I.15 Remplissage des bouteilles :

L'acétylène comprimé est transféré vers l'unité de remplissage où il est stocké dissous dans de l'acétone contenue dans une masse poreuse à l'intérieur des bouteilles. Le remplissage est effectué à une pression comprise entre 17 et 20 bars à 15 °C, garantissant la stabilité du gaz et la sécurité du stockage.



Figure I.10. Vue des réservoirs de stockage des gaz industriels au sein de Rayanox (Photo
Prise sur site de Rayanox, 2026)

I.15.1 Unité de remplissage et de mélange des gaz

Cette entité est chargée de l'emballage des gaz médicaux et industriels dans des cylindres haute pression ainsi que la conception des mélanges gazeux spécialisés.

C'est l'étape finale avant la livraison des produits aux clients du secteur industriel, hospitalier et agroalimentaire.

L'opération de remplissage est effectuée grâce à un système automatisé nommé « Techno Cryo Star », ce qui permet d'avoir un contrôle minutieux :

- Du flux
- De la pression
- De la température
- De la sûreté lors du remplissage.

Le procédé de remplissage comprend généralement

- Le raccordement des bouteilles au système
- Le double rinçage des bouteilles afin d'éliminer les contaminants
- La mise sous vide
- Le remplissage sous pression contrôlée
- Le contrôle final des paramètres de sécurité.



Figure I.11.Opération de remplissage et de conditionnement des bouteilles de gaz industriels dans l'unité de production de Rayanox (Rayanox, n.d.)

L'opération de remplissage des bouteilles d'oxygène médical se réalise à :

- Une pression d'environ 200 bars
- Une température comprise entre 15 et 20 °C.

L'unité comprend aussi : des systèmes de remplissage automatisés, des plateformes de contrôle et de commande, des équipements de sécurité contre les surpressions et les dangers d'incendie.

I.16 Politique Qualité, Hygiène, Sécurité et Environnement (QHSE)

Afin de garantir une production répondant aux normes industrielles, médicales et environnementales, l'entreprise adopte un système de gestion QHSE (Qualité, Hygiène, Sécurité et Environnement). Cet outil est un composant crucial dans l'arrangement et le pilotage des opérations industrielles de la société.

Rayanox applique un système de management QHSE visant (**Rayanox , s.d.**) :

- La qualité des produits ;

- La sécurité des travailleurs ;
- La protection de l'environnement ;
- L'amélioration continue.

L'entreprise possède plusieurs certifications :

- ISO 9001 : management de la qualité ;
- ISO 22000 : management des denrées alimentaires ;
- ISO 13485 : management de la qualité des dispositifs médicaux ;
- BPF : Bonnes Pratiques de Fabrication.



Figure I.12. Illustration du système QHSE intégrant la Qualité, l'Hygiène, la Sécurité et l'Environnement dans les milieux industriels (Tour de France des Alternatives, s.d.).

Conclusion

Pour résumer, RAYANOX s'est affirmé comme un acteur clé dans les domaines industriel et de la santé en Algérie. Cette réussite est due à son avancée technologique de premier plan, son offre de production variée (articles médicaux, industriels et alimentaires) et son protocole de certification QHSE strict. La société conjugue performance logistique et rigoureux respect des normes internationales afin de répondre de manière pérenne aux demandes des marchés domestiques et mondiaux.

Dans le chapitre suivant, nous aborderons l'intelligence artificielle, ses différentes catégories, ses principaux algorithmes ainsi que son rôle dans le domaine industriel.

Références Bibliographiques

- Algerie Eco.** (2021, 25 août). Industrie : L'usine Rayanox produira 100 000 litres d'oxygène par jour. *Algerie Eco*. <https://algerie-eco.com/2021/08/25/industrie-lusine-rayanox-produira-100-000-litres-doxygene-par-jour/>
- Boulenerouar, M.** (2021). Gestion de la maintenance préventive et sécurité des installations de gaz sous haute pression. *Revue des Sciences et Technologies Industrielles (USTO)*, 8(2), 88–102.
- Bouzidi, N.** (2018). Le développement des zones industrielles et leur impact sur l'attractivité territoriale : Cas de la zone industrielle de Bethioua (Oran). *Revue des Sciences Commerciales et Économiques*, 12(2), 45–60.
- Direction Européenne de la Qualité du Médicament & Soins de Santé.** (2020). *Pharmacopée Européenne* (10e éd.). Conseil de l'Europe.
- Haton, J.-P.** (2021, 10 août). Introduction to Artificial Intelligence. *Techniques de l'Ingénieur*.
- Kohl, A. L., & Nielsen, R. B.** (2017). *Gas Purification* (6e éd.). Gulf Professional Publishing.
- Messaoudi, A.** (2019). Analyse et contrôle de pureté des gaz industriels et médicaux par chromatographie. *Journal de la Société Algérienne de Chimie*, 29(1), 12–25.
- Radio Algérie.** (2021, 26 juillet). Covid-19 : Le ministère de l'Industrie annonce la réquisition de tous les moyens disponibles pour la production d'oxygène. *Radio Algérienne*. <https://www.radioalgerie.dz>
- SARL RAYANOX.** (s. d.). *Site officiel de RAYANOX*. <https://rayanox.co/>
- Tour de France des Alternatives.** (s. d.). *Illustration du système QHSE intégrant la Qualité, l'Hygiène, la Sécurité et l'Environnement dans les milieux industriels*. <https://alternatives.fr/>
- Zerrouki, M.** (s. d.). *Rapport de stage – Rayanox* [Rapport de stage non publié].

Chapitre II. L'intelligence artificielle et leurs applications dans le domaine industriel



II. Chapitre II. L'intelligence artificielle et leurs applications dans le domaine industriel

II.1 Introduction

L'industrie se caractérise par l'adoption généralisée des technologies digitales, parmi lesquelles l'intelligence artificielle (IA) se positionne comme le vecteur majeur de l'innovation (**Russell et Norvig, 2020**). En équipant les machines de compétences d'apprentissage autonome et de prise de décision, l'IA transcende la simple automatisation afin de rénover en profondeur les méthodes de production traditionnelles (**Schwab, 2017**). Dans ce chapitre, notre objectif est de cerner les notions essentielles de l'intelligence artificielle et d'examiner ses principales utilisations dans l'industrie. Nous prévoyons notamment d'étudier son influence sur la maintenance prédictive, le contrôle de qualité et l'amélioration des flux logistiques (**Lee et al., 2019**), dans le but de démontrer son véritable effet sur l'efficacité opérationnelle des usines contemporaines.

II.2 Définition de l'intelligence artificiel

C'est un processus qui reproduit l'intelligence humaine grâce à la conception et l'implémentation d'algorithmes fonctionnant dans un environnement informatique en évolution. Son objectif est de doter les ordinateurs de la capacité de penser et d'agir à l'instar des êtres humains.

Pour réaliser cela, il faut trois éléments :

- Des systèmes de technologie de l'information.
- Des informations relatives à des systèmes de gestion sont présentes.
- Des codes exploitant des algorithmes d'IA de pointe.

Afin de reproduire au mieux le comportement humain, l'intelligence artificielle exige un volume élevé de données et une puissance de traitement significative **NetApp.(s. d.)**.

Selon la définition de la norme ISO 2382-28, l'intelligence artificielle est la « capacité d'une unité fonctionnelle à réaliser des tâches généralement liées à l'intelligence humaine, comme le raisonnement et l'apprentissage ». Souvent qualifiée de prochaine révolution technologique, l'intelligence artificielle est au centre de tous les débats contemporains. Il semble crucial d'établir une définition pour cette technologie disruptive et de déterminer son cadre légal, en plus de repérer les applications existantes ou en développement au sein des entreprises et les avantages qu'elles en retirent (**CIGREF. 2016**).

II.3 Historique de l'IA

L'IA a une histoire qui s'étend sur plusieurs décennies. C'est dans les années 1950 que les chercheurs ont commencé à s'intéresser à ce domaine. De nombreuses méthodes ont été mises au point depuis, pour optimiser les facultés des machines en matière d'apprentissage et de raisonnement (CIGREF, 2016).

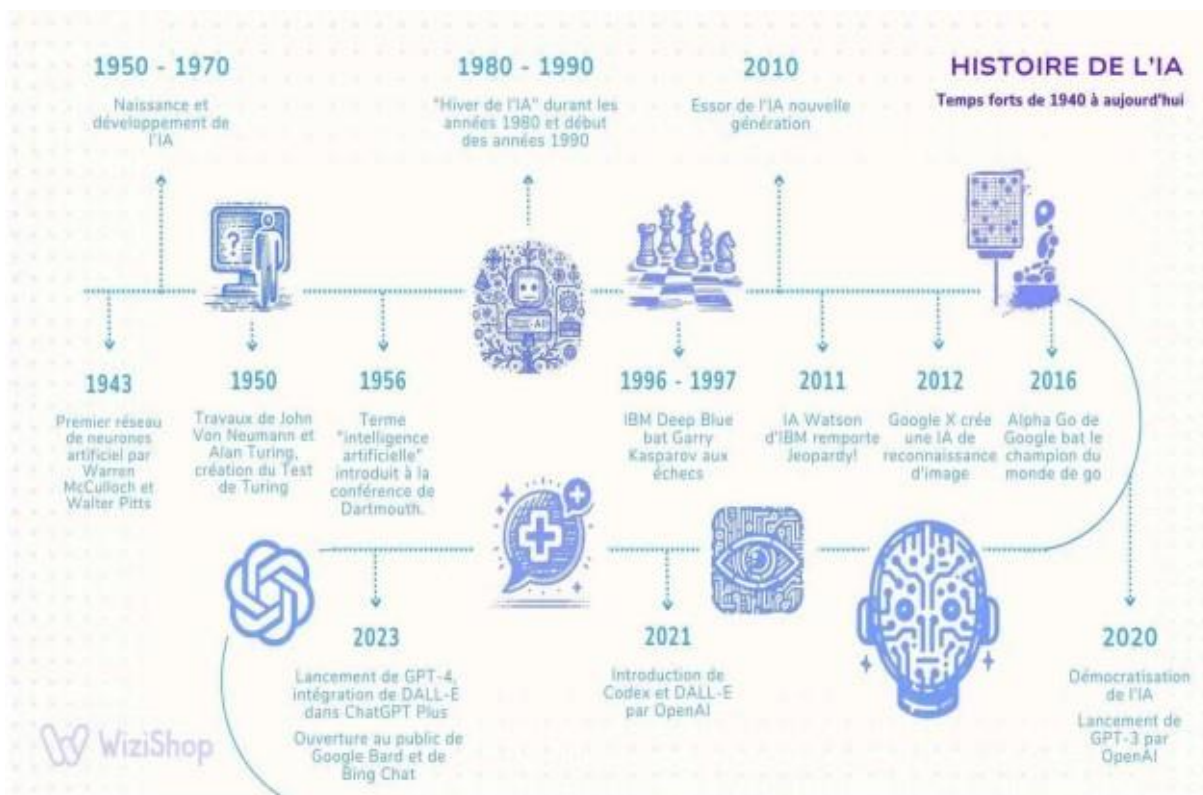


Figure II.1. Histoire de l'IA Beyrouti, G., & Cotton, R. (s. d.).

II.4 L'objectif de l'intelligence artificielle

Le but de l'intelligence artificielle est d'imiter le fonctionnement du cerveau humain en créant des « machines intelligentes ». Il est important de préciser qu'il s'agit non pas de machines dotées d'intelligence, mais bien de programmes faisant preuve d'intelligence. L'intelligence se manifeste de plusieurs façons : nous avons la capacité de développer notre savoir, de solutionner des problèmes, de maîtriser une langue, de retenir des données, et ainsi de suite (Hudault, B. 1988).

II.5 Les domaines d'application de l'IA

Selon **DataBird. (2026)**,

1. **Santé** : Diagnostic assisté par IA

L'IA contribue à l'interprétation des images médicales (comme les radiographies et les IRM) pour identifier les anomalies, facilitant ainsi des diagnostics plus rapides et exacts. Elle peut également contribuer à prédire des maladies en effectuant une analyse directe de données de santé.

2. **Finance** : Analyse prédictive et détection des fraudes

Il est possible d'utiliser des méthodes d'intelligence artificielle pour examiner des données financières en vue de prévoir les tendances de marché et les habitudes des consommateurs. On peut même repérer des actes frauduleux en décelant des comportements douteux dans les opérations financières.

Commerce : Personnalisation et chatbots

Les systèmes d'IA adaptent l'expérience utilisateur en examinant les goûts d'achat et en suggérant des produits appropriés. Grâce à l'IA, les chatbots permettent un service client efficace et immédiat.

Industrie : Maintenance prédictive et automatisation

L'intelligence artificielle est capable d'anticiper les défaillances d'appareils en examinant les données de performance, contribuant ainsi à minimiser les périodes d'inactivité. Une autre utilisation possible : l'automatisation des processus de production pour augmenter l'efficacité.

II.6 Les types d'intelligence artificielle

Il y a un consensus parmi les chercheurs pour distinguer trois types majeurs d'intelligence artificielle, basés sur leur niveau de compétence et leur niveau d'autonomie cognitive (**Liora, s. d.**).

Tableau II.1: Les types d'intelligence artificielle

Type d'IA	Niveau de capacité	Statut actuel	Exemple
IA faible (Narrow AI)	Tâche unique et spécifique	Existante et très répandue	Recommandations Netflix, reconnaissance vocale
IA forte	Conscience et sentiments propres	Théorique, inexistante	Scénarios de science-fiction
IA générale (AGI) et super-intelligence	Toutes les tâches cognitives humaines, puis au-delà	Hypothétique, en recherche active	AGI : débat autour de GPT-4 ; Super-IA : concept futur

II.6.1 IA faible (Narrow AI)

L'intelligence artificielle faible, également connue sous le nom d'intelligence étroite ou Narrow AI, est un système élaboré pour accomplir une tâche spécifique de façon pratiquement impeccable. C'est le genre d'IA le plus couramment employé jusqu'à présent. Des exemples incluent les assistants vocaux tels que Siri et Alexa, les systèmes de recommandation de Netflix et Spotify, la technologie de reconnaissance faciale, les filtres anti-spam, ainsi que les chatbots utilisés pour le service client. En dépit de ses bénéfices, des données d'entraînement de piètre qualité peuvent causer des résultats déformés aux implications sévères dans des domaines délicats tels que l'autorisation de crédits ou l'embauche (**Liora, s. d.**).

II.6.2 IA forte

On évoque l'Intelligence Artificielle forte quand un modèle est théoriquement considéré comme possédant une existence intérieure authentique : des émotions, de la subjectivité, une conscience de soi. Au sujet de l'intelligence artificielle forte, qui semblerait tout droit sortie d'un récit de science-fiction, les spécialistes estiment qu'elle ne peut pas être réalisée actuellement : l'idée de conscience ne peut émerger dans des dispositifs qui ne font que traiter des symboles et des calculs (**Liora, s. d.**).

II.6.3 IA générale (AGI) et super-intelligence

L'AGI (Intelligence Artificielle Générale) est une intelligence artificielle qui peut accomplir tout type de tâche cognitive à l'instar d'un être humain. Étant encore au stade hypothétique, certains chercheurs se questionnent sur GPT-4 en tant que première manifestation d'une AGI. Au-delà, on désigne par super-intelligence artificielle une entité dotée de conscience de soi, autonome par rapport au contrôle humain et nettement plus intelligente que l'homme. Ces idées provoquent de fortes préoccupations éthiques, ce qui rend nécessaires des encadrements comme l'AI Act en Europe (**Liora, s. d.**).

II.7 Les principaux modèles et architectures d'IA

Il existe plusieurs catégories de modèles d'intelligence artificielle, chacune étant basée sur des architectures et des méthodes spécifiques. De l'apprentissage automatique traditionnel aux grands modèles multimodaux, voici une vue d'ensemble des architectures clés qui sculptent l'intelligence artificielle contemporaine (**Liora, s. d.**).

II.7.1 Le machine Learning ou apprentissage automatique

Le Machine Learning, branche de l'intelligence artificielle, est l'aspect de l'informatique qui vise à analyser et à décoder des motifs et agencements de données en vue de faciliter l'apprentissage, le raisonnement et les décisions indépendamment de toute interaction humaine. En synthèse, grâce à l'apprentissage automatique, l'utilisateur fournit à un algorithme une quantité importante de données, et demande à la machine de formuler des suggestions et de faire des choix fondés sur ces données. Dans le cas où des rectifications sont effectuées, l'algorithme peut intégrer ces données pour optimiser ses décisions.

Ainsi, on vise à l'utilisation du Machine Learning pour élaborer des systèmes capables d'apprendre et de progresser à partir de l'expérience, au lieu d'être configurés avec des directives précises pour chaque mission (**LARIBI et al., 2023**).

II.7.2 Le fonctionnement de la machine Learning

Le Machine Learning se compose de trois volets distincts :

- L'algorithme qui fait partie intégrante du processus de prise de décision.
- Les facteurs et les caractéristiques qui déterminent la décision.

Chapitre II. L'intelligence artificielle et leurs applications en industrie

- La base de connaissances grâce à laquelle la réponse est connue, et qui permet au système d'évoluer.

Dès le début, le modèle est nourri avec des données de paramètres issues de réponses déjà connues. L'algorithme est par la suite exécuté et des modifications sont apportées jusqu'à ce que le dénouement de celui-ci (l'apprentissage) corresponde à la réponse déjà connue. Actuellement, on introduit une quantité toujours croissante de données afin de former le système pour gérer des décisions de plus en plus complexes (**LARIBI et al., 2023**).

II.8 Les différents types de Machine Learning

Habituellement, on catégorise les techniques d'apprentissage automatique selon la quantité d'effort requis.

Un humain est indispensable pour les coordonner et pour la manière dont elles exploitent les données marquées.

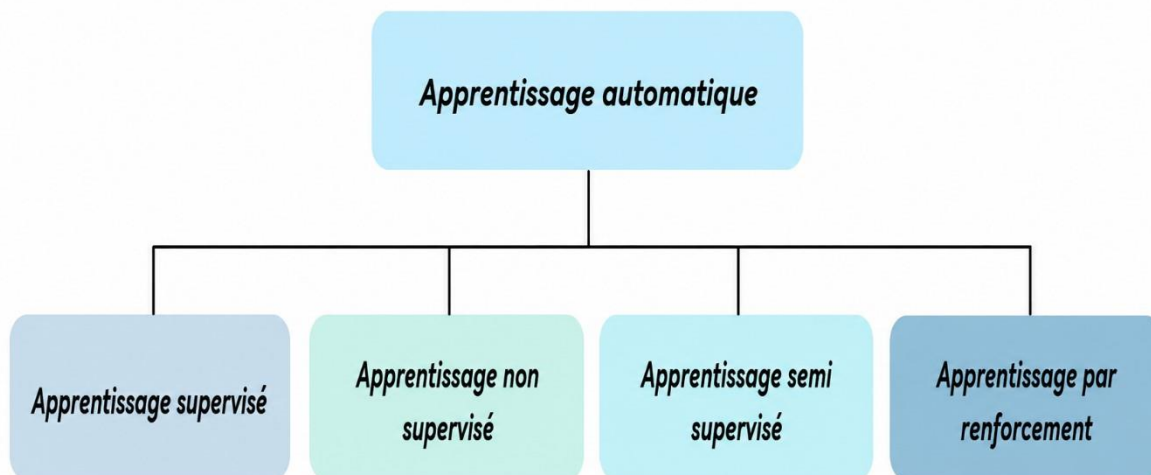


Figure II.2. Types de Machine Learning. ResearchGate GmbH (2024)

II.8.1 Apprentissage Supervisé

Dans le cadre de l'apprentissage supervisé, un algorithme est formé pour effectuer des prédictions ou des choix en s'appuyant sur des données d'apprentissage étiquetées. En d'autres termes, les données d'entrée sont liées à une étiquette de sortie ou à une valeur cible associée, et l'algorithme se perfectionne à établir la correspondance entre l'entrée et la sortie.

Chapitre II. L'intelligence artificielle et leurs applications en industrie

L'apprentissage supervisé peut être administré en deux sections distinctes : la régression et la classification (**LARIBI et al., 2023**).

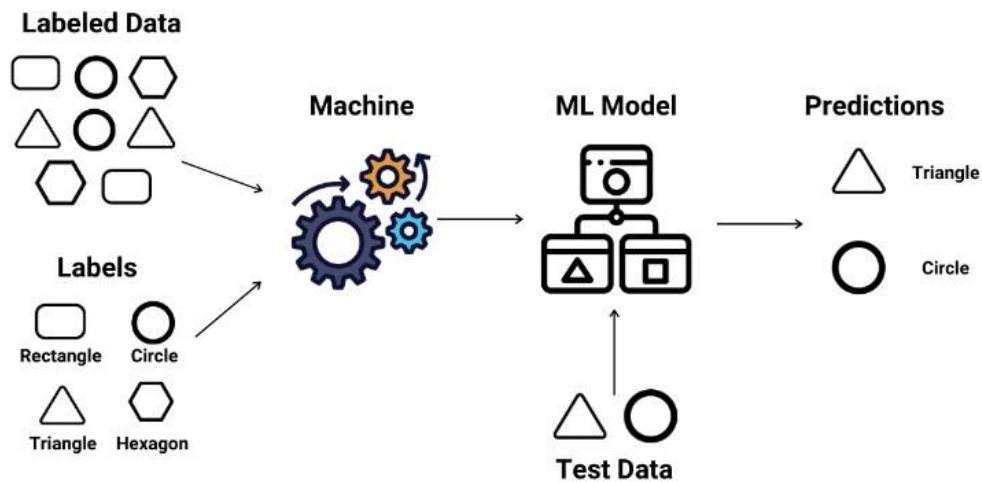


Figure II.3. Apprentissage Supervisé. Blent (2026).

II.8.2 Régression linéaire

Modèle statistique facilite l'établissement d'un lien prédictif clair entre une variable continue (Qu'on appelle dépendante ou « expliquée ») et une ou plusieurs variables indépendantes (qu'on nomme « explicatives »). L'évaluation des coefficients mathématiques par la méthode des moindres carrés est effectuée par le modèle dans le but de réduire l'écart quadratique entre les valeurs observées et les valeurs anticipées (**Houde, s. d.**). Dans le domaine de l'industrie, elle est essentielle pour établir un lien direct entre les variables physiques globales.

II.8.3 Régression Linéaire Simple

La régression linéaire simple constitue une technique statistique fondamentale en apprentissage supervisé. Elle est utilisée pour modéliser et analyser la relation linéaire entre deux variables continues : une variable explicative (ou indépendante), notée X, et une variable dépendante (ou variable à expliquer), notée Y (**Zou et al., 2003**).

Cet algorithme vise principalement à établir la « droite de meilleur ajustement » (best-fit line) à l'aide de la technique des moindres carrés, qui cherche à réduire la somme des carrés des différences (résidus) entre les données réelles et les données prévues par le dispositif (**Zou et al., 2003**).

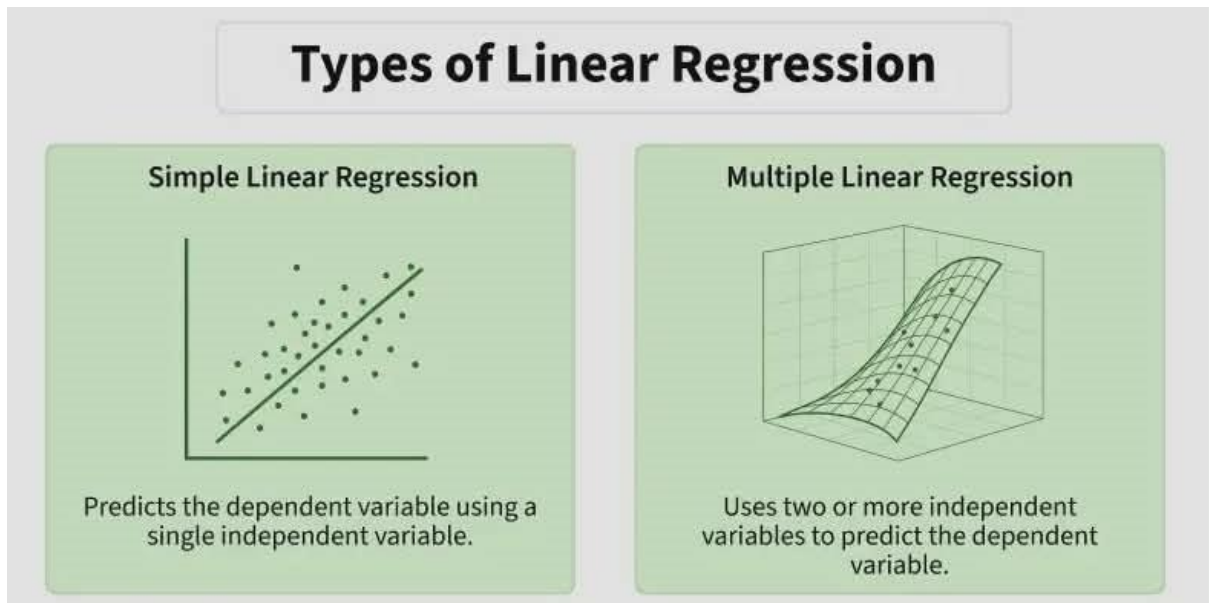


Figure II.4. Différents modèles de régression linéaire simple vs multiple

Le modèle mathématique s'exprime sous la forme suivante :

$$Y = \beta_0 + \beta_1 x + \varepsilon \quad \text{II.1}$$

Où :

β_0 : représente l'ordonnée à l'origine (intercept).

β_1 : est la pente de la droite de régression (coefficient).

ε : est le terme d'erreur résiduelle.

II.8.4 Régression Linéaire Multivariable

La régression linéaire multivariable (également connue sous le nom de régression linéaire multiple) constitue une version développée de la régression linéaire simple. Elle sert à représenter le lien entre une variable dépendante continue y (par exemple, la consommation d'énergie ou de GPL) et plusieurs variables indépendantes ou prédictives $x_1, x_2, \dots, x_k$ (telles que le débit, la pression ou la température) (Chen, 2026).

Le modèle mathématique général de la régression linéaire multiple s'exprime sous la forme suivante :

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad \text{II.2}$$

Où :

y : est la réponse ou la variable dépendante.

Chapitre II. L'intelligence artificielle et leurs applications en industrie

$x_1, \dots, x_k$: sont les variables prédictives (indépendantes).

$\beta_0, \beta_1, \dots, \beta_k$: coefficients

ε : représente le terme d'erreur aléatoire.

Les métriques MAE, RMSE et R^2 sont indispensables pour évaluer les performances d'un modèle de régression. Elles permettent de mesurer l'erreur de prédiction et la qualité d'ajustement global du modèle par rapport à vos données réelles.

II.8.5 MAE (Mean Absolute Error - Erreur Absolue Moyenne)

La MAE calcule la moyenne des écarts absolus entre les valeurs réelles et les valeurs prédites. Elle est très robuste face aux valeurs aberrantes.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad \text{II.3}$$

$\hat{Y}_i$: valeur prédite;

Y_i : valeur réelle ;

n : nombre total d'observation.

II.8.6 RMSE (Root Mean Squared Error - Erreur Quadratique Moyenne)

La RMSE est la racine carrée de la moyenne des erreurs élevées au carré. Le fait d'élever les erreurs au carré pénalise beaucoup plus les fortes erreurs (valeurs aberrantes) que la MAE.

$$RMSE = \sqrt{\sum_{i=1}^n \frac{1}{n} (y_i - \hat{y}_i)^2} \quad \text{II.4}$$

. Arbre de Décision (Decision Tree)

Cet algorithme emploie une procédure de segmentation récursive de l'espace de données afin d'élaborer une structure arborescente hiérarchisée. À chaque branche, une évaluation est réalisée sur un critère opérationnel pour séparer les données en sous-groupes de plus en plus pures. Pour les tâches de régression, la sélection de la coupe idéale est basée sur la réduction de la variance résiduelle des segments descendants (**Decaestecker & Saerens, n.d.**).

II.8.7 Forêt Aléatoire (Random Forest)

Cet algorithme, élaboré comme une version avancée des arbres de décision, s'appuie sur l'apprentissage collectif (Ensemble Learning). Il repose sur la création d'un vaste ensemble d'arbres de décision autonomes, formés sur des sous-groupes de données grâce à la méthode du bagging, couplée à une sélection aléatoire des variables à chaque nœud (**Breiman, 2001**).

Exemple : Estimer la consommation de GPL (en tonnes) du jour suivant en tenant compte de la température extérieure projetée et de la quantité de CO₂ à émettre.

II.8.8 Algorithme XGBoost

L'algorithme XGBoost (Extreme Gradient Boosting) s'affirme comme une des méthodes d'ensemble les plus redoutables et efficaces de l'apprentissage supervisé, bénéficiant du renforcement des performances des arbres de décision et du Random Forest. Cet algorithme, qui repose sur le concept du « Gradient Boosting », ne forme pas les arbres de façon autonome, mais plutôt successive : chaque arbre nouvellement créé est précisément conçu pour rectifier les erreurs résiduelles laissées par les arbres antérieurs (**Mesut & Aksu, 2023**).

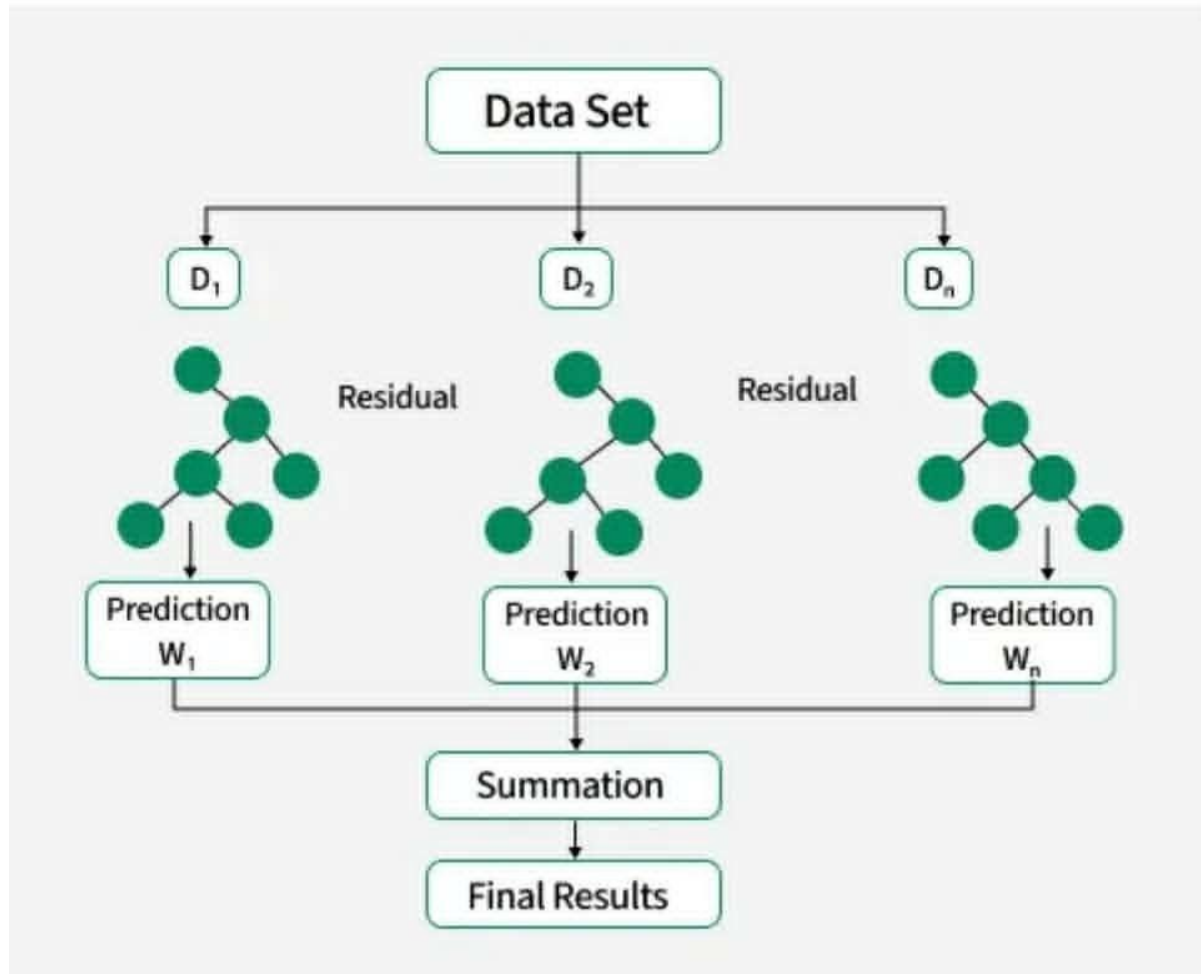


Figure II.5. Architecture séquentielle de l'algorithme XGBoost et processus de correction des erreurs résiduelles. (Adapté de GeeksforGeeks, 2025).

II.8.9 Apprentissage non Supervisé

Le but de l'apprentissage non supervisé est de détecter des tendances ou des configurations dans les informations, sans utiliser d'étiquettes ou de buts définis. Autrement dit, les données que vous manipulez ne sont pas étiquetées, et vous devez permettre à l'algorithme de déduire de manière indépendante la structure sous-jacente et les liens entre ces données (**LARIBI et al., 2023**).

Apprentissage non supervisé

L'algorithme découvre des structures ou des groupes dans des données non étiquetées.

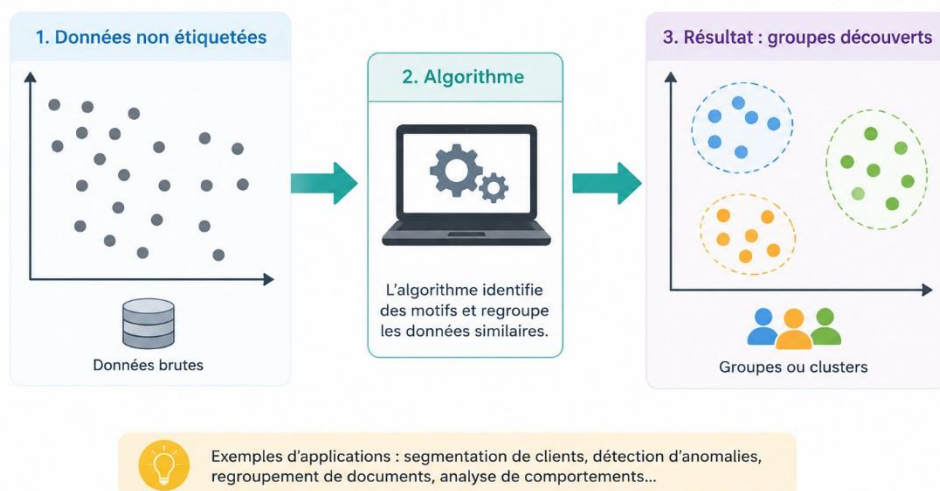


Figure II.6. Apprentissage non Supervisé (Réalisation personnelle)

II.8.10 Apprentissage semi-supervisé

La méthode de l'apprentissage semi-supervisé combine des éléments des deux types précédents. Dans ce genre de formation, un groupe restreint de données labellisées est associé à un lot plus conséquent de données non labellisées pour former un modèle. Les données annotées sont utilisées comme référence durant le processus d'apprentissage, tandis que l'utilisation de données non annotées permet au modèle de se généraliser et d'obtenir des caractéristiques plus solides. Cette méthode est particulièrement utile dans les contextes où l'acquisition d'une vaste quantité de données labellisées peut s'avérer complexe ou onéreuse (LARIBI et *al.*, 2023).

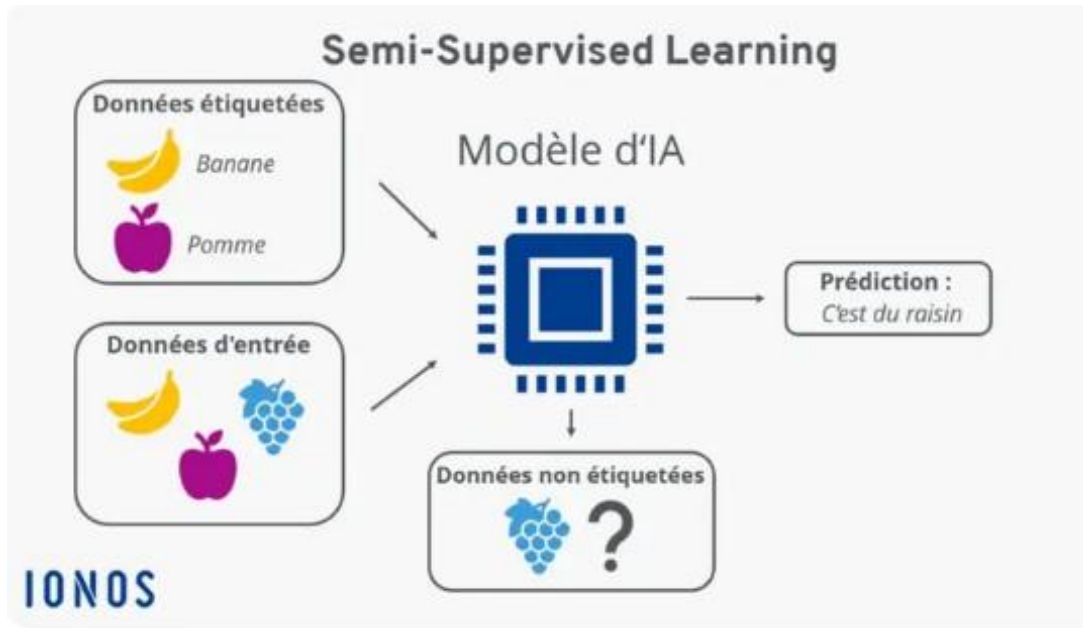


Figure II.7. Apprentissage semi-supervisé (IONOS ,2026).

II.8.11 L'apprentissage par renforcement

Est une stratégie qui consiste pour un agent à interagir avec un environnement dans le but d'apprendre à effectuer des actions qui maximisent une récompense cumulative.

L'agent progresse en apprenant par l'expérience, recevant des rétroactions sous la forme de gains ou de sanctions en lien avec ses actions. Cette méthode se base sur trois composantes centrales : l'agent, l'environnement et le signal de récompense. L'agent agit sur la base de son état présent en environnement et reçoit un retour de récompense indiquant la justesse ou l'erreur de ces actions. L'agent exploite les retours qu'il a reçus pour réviser sa stratégie et perfectionner ses mesures à venir (LARIBI *et al.*, 2023).

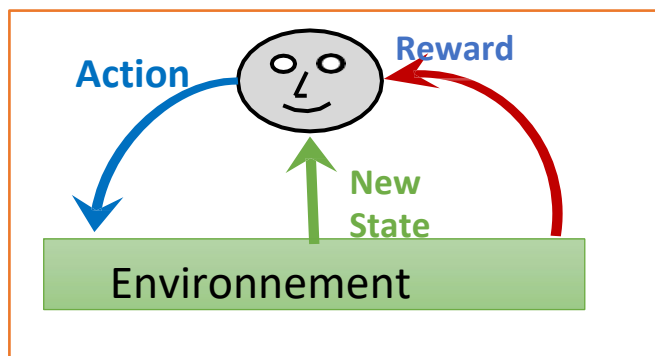


Figure II.8. Apprentissage par renforcement.

Chapitre II. L'intelligence artificielle et leurs applications en industrie

II.8.12 Deep Learning

Le Deep Learning représente une évolution sophistiquée du Machine Learning, exploitant des réseaux neuronaux profonds pour gérer d'énormes volumes de données. À l'opposé du Machine Learning traditionnel qui requiert fréquemment une identification manuelle des caractéristiques, le Deep Learning est capable d'apprendre de manière indépendante des représentations hiérarchiques directement à partir des données brutes, à titre d'exemple un modèle peut de ce fait différencier un chat d'un chien sur une image sans qu'un développeur n'ait précisé ce qu'est une oreille ou un museau (**Liora. s. d.**).

II.9 LLM et modèles multimodaux

Les Modèles de Langage de Grande Taille (LLM) sont des réseaux de neurones de type Transformer formés sur d'énormes volumes de données textuelles, dotés de la capacité à comprendre et à produire du langage naturel de façon extrêmement fluide. Le phénomène le plus marquant est l'apparition des modèles multimodaux, aptes à gérer en même temps le texte, l'image, l'audio et la vidéo. Des modèles tels que GPT-4o ou Gemini sont capables d'examiner une image, de répondre à des interrogations à partir d'un document numérisé ou de transcrire une bande sonore - ils traitent et produisent de l'information de manière aussi variée qu'un être humain (**Liora. s. d.**).

II.10 Comment entraîner et déployer une intelligence artificielle

L'élaboration d'une intelligence artificielle efficace englobe toutes les étapes, depuis la collecte de données jusqu'à la surveillance en environnement opérationnel :

- ✓ Collecte et préparation des données

La qualité du modèle dépend directement de la qualité des données : un ensemble de données biaisé ou incomplet entraînera assurément des prédictions erronées. Voici les étapes clés :

- ✓ Détermination des sources : bases de données internes, interfaces de programmation d'applications, fichiers à plat, données issues du web ou capteurs.
- ✓ Nettoyage et dédoublonnage : élimination des données manquantes, rectification des irrégularités et uniformisation des formats.
- ✓ Taggage : dans le cadre de l'apprentissage supervisé, chaque donnée doit être attribuée à une réponse prévue.
- ✓ **Feature engineering** : choix et modification des variables significatives.

Division en ensembles d'entraînement, de validation et de test : segmentation du jeu de données pour former, optimiser et juger le modèle de manière autonome

(Liora. s. d.).

✓ Entraînement et optimisation du modèle

Le modèle est alimenté par les données afin de déceler des motifs. Le modèle étudie les erreurs de prédiction grâce aux fonctions de coût, puis ajuste ses poids internes en utilisant la technique de descente de gradient. L'amélioration implique particulièrement l'ajustement précis des hyperparamètres, taux d'apprentissage, taille du lot, nombre d'époques, qui ont un impact direct sur la capacité à généraliser et à prévenir le surapprentissage. Des méthodes telles que la validation croisée k-fold, le search par grille et l'optimisation bayésienne sont utilisées pour définir l'ajustement optimal des paramètres **(Liora. s. d.)**.

✓ Déploiement, MLOps et monitoring

Un modèle correctement formé ne génère véritablement de la valeur que lorsqu'il est mis en production. La notion de MLOps (Machine Learning Operations) englobe toutes les méthodes visant à rendre industrielles, automatiques et supervisées les modèles en production. Une fois mis en œuvre par le biais d'une API, il est essentiel de superviser constamment le modèle pour repérer tout data drift, la déviation graduelle des données réelles par rapport à celles utilisées lors de l'entraînement, qui pourrait, avec le temps, détériorer les performances de manière insidieuse. Un pipeline MLOps efficace assure une reproductibilité, une collaboration et une stabilité adéquate en production **(Liora. s. d.)**.

II.11 Comment évaluer la performance d'une IA

Selon **(Liora. s. d.)**, les chercheurs et les ingénieurs utilisent des indicateurs chiffrés, appelés métriques, pour évaluer objectivement les capacités d'une IA. Ils se servent également de benchmarks, des ensembles de tests standardisés, pour comparer les performances des modèles sur des tâches spécifiques.

Tableau II.2.Les métriques clés

Métrique	Ce qu'elle mesure		Quand l'utiliser
Accuracy	Pourcentage de prédictions correctes		Classification de texte, réponse à des questions
Score F1	Équilibre entre précision et rappel		Données déséquilibrées, prédictions catégorielles
Perplexité	Capacité d'un LLM à prédire un échantillon de texte		Évaluation des grands modèles de langage
Score BLEU	Qualité de la traduction entre deux langues		Tâches de traduction automatique
Fonction de coût (MSE ou RMSE)	Mesure l'erreur entre les valeurs réelles et prédites		Modèles de régression
Coefficient de détermination (R²)	Mesure la qualité d'ajustement du modèle		Régression linéaire et prévision
MMLU	Connaissances générales et raisonnement (LLM)	16 000 questions à choix multiples couvrant 57 disciplines académiques, testant mémorisation factuelle et raisonnement. Incontournable dans les rapports techniques sur les LLM.	

II.12 L'IA enjeu pour l'industrie

L'IA, principalement fondée sur les technologies de base comme les réseaux neuronaux et le deep learning, suscite aujourd'hui un intérêt immense. La force des ordinateurs dissimule l'aspect complexe des traitements associés à ces technologies. L'intelligence artificielle a gagné des domaines d'application diversifiés : éducation, finance, hôpitaux, médecine, recrutement, service à la clientèle, jeux, maintenance, parmi d'autres.

C'est indubitable que les domaines les plus prometteurs sont ceux qui se rapportent à la reconnaissance des formes et à la gestion de la parole. On utilise fréquemment les systèmes de lecture optique, tout comme la technologie de reconnaissance et de synthèse vocale.

Chapitre II. L'intelligence artificielle et leurs applications en industrie

Les assistants « intelligents » se sont donc rendus courants, et comme mentionné plus tôt, ils laissent pour l'instant une forte impression sur le public **Hauet, J.-P. (2018)**.

Cependant, l'IA a également dominé le secteur professionnel, notamment en médecine, où l'analyse d'images ou de radiographies peut désormais être efficacement déléguée à des machines qui exploitent l'IA.

Il est donc logique de se demander quel sera le rôle de l'IA dans l'industrie à venir et quelle aide elle pourrait offrir à la revitalisation industrielle du pays **(Hauet, J.-P. 2018)**.

Aujourd'hui, les domaines où l'IA peut être appliquée dans l'industrie se sont grandement étendus. Une des justifications de cela est l'évolution avancée des méthodes d'IA qui rend envisageable leur transposition des domaines grand public vers les domaines professionnels, à l'image de ce qui a eu lieu pour les applications liées à la médecine. Une autre raison concerne l'explosion des données avec lesquelles le secteur industriel sera amené à travailler et qu'il devra gérer. Selon McKinsey, depuis 2010, le secteur industriel est le principal générateur de données, produisant le double de données par rapport aux gouvernements, institutions bancaires, communications & médias, et plus de quatre fois la quantité de données produite dans la distribution. L'avancement des capteurs intelligents et communicants, essentiels à l'IIoT (Internet Industriel des Objets), rend désormais accessible une multitude de données plus détaillées et diversifiées. L'utilisation de méthodes reposant sur l'IA est actuellement souhaitée pour exploiter ces énormes volumes de données. Comme le notait la REE dans son analyse IoT 2018, nous sommes passés de l'époque de la gestion des données à celle de la gestion par les données. L'information prend le pas et devient le véritable moteur de l'économie contemporaine **(Hauet, J.-P. 2018)**.

Tableau II.3 .Tableau des domaines d'application potentiels de l'IA.

	Domaines	Cas d'usage typiques
Extension de techniques grand public	Reconnaissance de formes et de caractère - Vision	• Contrôles de fabrication • Détection de défauts
	Traitement du langage naturel - Reconnaissance de la parole - Synthèse vocale	• Interface clients • Assistants personnels pour équipes d'intervention et de maintenance • Formation des personnels • Automatisation de la création et de la mise à jour de la documentation
Technologies professionnelles	Organisation et contrôle des procédés	• Robotique et automatisation des procédés • Optimisation de la conduite des procédés • Organisation de la maintenance prédictive • Organisation de la logistique – Gestion des approvisionnements et des entrepôts
	Analyse stratégique et aide à la décision	• Analyse des marchés, veille stratégique et concurrentielle, détection des signaux faibles • Affectation optimale des ressources, planification et gestion des systèmes complexes • Analyse prédictive
	Sécurité	• Contrôle d'accès • Cybersécurité : détection d'intrusions

Ce tableau 2.3 offre une catégorisation des secteurs d'application où l'intelligence artificielle peut actuellement s'insérer dans l'industrie. On différencie les applications qui s'inscrivent dans la continuité de celles qui sont déjà bien établies auprès du grand public, et celles qui sont plus orientées vers le secteur industriel.

L'intégration de l'IA dans l'industrie n'en est qu'à ses premiers pas et de nombreux efforts sont nécessaires pour inciter le milieu industriel à adopter ces technologies. Cependant, la question est cruciale. Dans un contexte de nécessité de personnalisation et d'évolution constante, face à la compétition accrue des pays en développement et aux attentes sans cesse croissantes des consommateurs, tant en matière de variété des produits que de qualité, l'industrie doit se transformer. Le rapport de Cédric Villani de mars 2018, intitulé « Donner un sens à l'intelligence artificielle », propose des orientations pour remettre la France parmi les leaders de l'IA. Positionnée en plein cœur de la révolution industrielle, l'intelligence artificielle repousse les frontières de l'automatisation et de l'efficacité (gdd, gdsg, p. 12)(Hauet, J.-P. 2018).

II.13 Travaux antérieurs de l'application de l'intelligence artificielle dans le domaine industrielle

L'intelligence artificielle (IA) a gagné en popularité dans le secteur industriel, notamment grâce à ses applications variées dans l'optimisation des processus, la maintenance prédictive, le contrôle de qualité, et la robotique. Voici principales études et applications de l'IA dans le domaine industriel :

a) Optimisation des Processus

- **Étude sur la planification intelligente :** De nombreux travaux ont exploré l'utilisation de l'apprentissage automatique pour optimiser la planification de la production. Par exemple, Wang et al. (2020) ont développé un algorithme basé sur des réseaux de neurones pour améliorer l'efficacité de la chaîne d'approvisionnement dans l'industrie automobile.
- **Simulation et modèles prédictifs :** Les modèles de simulation intégrant l'IA, comme ceux décrits par Zhang et al. (2021), permettent de prévoir les performances des systèmes de production dans différents scénarios.

b) Maintenance Prédictive

- **Détection des anomalies :** Des applications basées sur le machine learning, comme celles de Ghanbari et al. (2018), utilisent des données de capteurs pour prédire les défaillances d'équipements, réduisant ainsi le temps d'arrêt.
- **Modèles de survie :** Les travaux de Babu et al. (2019) sur l'analyse des facteurs de risque ont montré comment l'IA peut être utilisée pour modéliser la durée de vie des machines industrielles.

c) Contrôle de Qualité

- **Vision par ordinateur :** Plusieurs études, comme celles d'Abdelrahman et al. (2020), appliquent la vision par ordinateur pour l'inspection de produits afin de détecter les défauts en temps réel, améliorant ainsi la qualité globale.
- **Analyse des données :** L'utilisation d'algorithmes d'apprentissage supervisé pour analyser les données historiques des défauts a été explorée par plusieurs chercheurs, permettant d'identifier des tendances et d'optimiser les processus.

d) Robotique et Automatisation

- **Robots autonomes** : Des chercheurs, tels que Kim et al. (2021), ont développé des systèmes robotiques autonomes utilisant des techniques d'IA pour exécuter des tâches complexes dans des environnements industriels.
- **Interaction homme-robot** : Des études ont montré l'importance d'intégrer l'IA dans les systèmes de collaboration homme-robot, ce qui permet une meilleure prise de décision et une amélioration de l'efficacité, comme discuté dans les travaux de Rojas et al. (2022).

e) Supply Chain Management

- **Prévisions de la demande** : L'IA est utilisée pour améliorer les prévisions de la demande dans les chaînes d'approvisionnement, comme l'ont montré les études de Chae (2019), qui ont intégré des algorithmes d'apprentissage profond pour anticiper les fluctuations du marché.
- **Optimisation des stocks** : L'utilisation d'analyses prédictives pour la gestion des stocks a été traitée dans plusieurs études, permettant de réduire les coûts et d'améliorer la disponibilité des produits.

Conclusion

L'intelligence artificielle transforme de nombreux aspects de l'industrie. Les études montrent son potentiel dans l'optimisation des performances, la réduction des coûts et l'amélioration de la qualité. Les recherches futures devraient se concentrer sur l'intégration de l'IA avec d'autres technologies émergentes comme l'Internet des objets (IoT) et la blockchain pour renforcer davantage l'efficacité des systèmes industriels.

Références Bibliographiques

- Abdelrahman, M., et al.** (2020). Real-time defect detection using computer vision.
- Babu, A., et al.** (2019). Survival analysis for equipment lifespan prediction.
- Beyrouti, G., & Cotton, R.** (s. d.). *Histoire de l'IA et ses jalons technologiques*. Tech Report.
- Blent.** (2026). *Introduction à l'apprentissage supervisé et ses algorithmes*. Blent AI Tech.
- Breiman, L.** (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Chae, B.** (2019). Demand forecasting using deep learning in supply chains.
- Chen, G.** (2026). *Multiple Linear Regression* [Lecture notes]. Department of Mathematics & Statistics, San José State University.
- CIGREF.** (2016). *Gouvernance de l'intelligence artificielle dans les grandes entreprises*. Rapport du Cigref.
- DataBird.** (2026). *Les grands domaines d'application de l'intelligence artificielle*. DataBird Institute.
- Decaestecker, C., & Saerens, M.** (s. d.). *Les arbres de décision (decision trees)* [Support de cours]. Université Libre de Bruxelles & Université Catholique de Louvain, Belgique.
- GeeksforGeeks.** (2025). *XGBoost algorithm: Architecture and residual error correction process*. GeeksforGeeks.
- GeeksforGeeks.** (2026). *Types of linear regression in machine learning*. GeeksforGeeks.
- Ghanbari, A., et al.** (2018). Predictive maintenance using machine learning.
- Hauet, J.-P.** (2018). L'intelligence artificielle: Quel rôle pour l'industrie? *Revue de l'Électricité et de l'Électronique (REE)*, (4), 34–45.
- Houde, L.** (s. d.). *Module 13: Régression linéaire simple (Analyse quantitative de problèmes de gestion - PAF-1010)* [Support de cours]. Département de Mathématiques et d'Informatique, Université du Québec à Trois-Rivières, Canada.
- Hudault, B.** (1988). *L'intelligence artificielle*. Presses Universitaires de France (PUF).
- International Organization for Standardization.** (1987). *Information processing systems — Vocabulary — Part 28: Artificial intelligence* (ISO Standard No. 2382-28:1987).

- IONOS.** (2026). *Qu'est-ce que l'apprentissage semi-supervisé?*. IONOS Digital Guide.
- Kim, J., et al.** (2021). Autonomous robotics in industrial applications.
- Laribi, A., et al.** (2023). *L'apprentissage automatique et ses applications industrielles* [Rapport technique]. Université de technologie.
- Lee, J., et al.** (2019). Industrial Artificial Intelligence for smart manufacturing systems. *Manufacturing Letters*, 18, 20–25.
- Liora.** (s. d.). *Guide complet sur les architectures d'intelligence artificielle, MLOps et benchmarks*. Liora AI Blog.
- Mesut, B., & Aksu, N. B.** (2023). Role of artificial intelligence in quality profiling and optimization of drug products. In *A Handbook of Artificial Intelligence in Drug Delivery* (pp. 145-170). Academic Press.
- NetApp.** (s. d.). *L'importance des données massives et du calcul en intelligence artificielle*. NetApp Tech Insights.
- ResearchGate GmbH.** (2024). *Classification and types of machine learning algorithms* [Diagram]. ResearchGate.
- Rojas, A., et al.** (2022). Human-robot collaboration in production systems.
- Russell, S., & Norvig, P.** (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Schwab, K.** (2017). *The fourth industrial revolution*. Portfolio Penguin.
- Villani, C.** (2018). *Donner un sens à l'intelligence artificielle: Pour une stratégie nationale et européenne*. Rapport parlementaire, Mission parlementaire sur l'intelligence artificielle.
- Wang, Y., et al.** (2020). Intelligent production planning using neural networks.
- Zhang, X., et al.** (2021). Simulation-based optimization in industrial processes.
- Zou, K. H., Tuncali, K., & Silverman, S. G.** (2003). Correlation and simple linear regression. *Radiology*, 227(3), 617-628.

CHAPITRE III. Résultats et discussion de la Modélisation de la consommation énergétique de l'unité de la production de CO₂



III.CHAPITRE III. Résultats et discussions de la modélisation de la consommation énergétique de l'unité de la production de CO₂

III.1 Introduction

Dans le cadre de ce stage effectué au sein de l'entreprise **Rayanox**, nous avons suivi de près le processus de collecte et d'exploitation des données relatives à la consommation quotidienne d'énergie électrique liée à l'unité de la génération de gaz carboniques. L'objectif principal de cette étude est d'analyser l'influence de plusieurs variables explicatives, à savoir la consommation totale de GPL, le total de CO₂ et la pression, sur la consommation d'électricité, afin de construire un modèle de prédiction fiable.

Pour répondre à cette problématique, nous avons choisi d'adopter une méthode basée sur l'apprentissage automatique (machine Learning). À cette fin, nous avons sélectionné le langage de programmation Python pour coder et interpréter nos algorithmes. Plusieurs algorithmes de régression ont été testés afin de comparer leurs performances et d'identifier le modèle le plus pertinent pour la prévision de la consommation énergétique. Les trois méthodes retenues sont la régression linéaire, le Random Forest et XG Boost, choisies pour leur complémentarité entre simplicité, robustesse et capacité de modélisation des relations complexes.

La présente partie consacrée aux résultats et à la discussion a pour but de présenter les performances obtenues par chacun de ces modèles, d'interpréter les indicateurs d'évaluation utilisés, et de comparer leurs capacités prédictives. Cette analyse permettra également de mettre en évidence la relation entre les variables étudiées et la consommation d'électricité, tout en discutant les avantages et les limites de chaque algorithme dans le contexte industriel de l'entreprise.

III.2 Collecte des données

La consommation électrique représente une part importante du coût de production du CO₂. Une prédiction fiable permet :

- Réduction des coûts énergétiques
- Optimisation de la production
- Amélioration de la rentabilité
- Aide à la prise de décision

L'information a été recueillie sur une période de deux mois, du 1er avril 2026 au 30 mai 2026. Malgré l'importance d'une base de données pour la modélisation par l'IA, nous n'avons réussi

CHAPITRE III. Résultats et discussion de la modélisation de consommation énergétique de l'unité de la production de CO₂

à accumuler que 60 ensembles de données. Pour mettre en œuvre la prédiction des données grâce à l'apprentissage automatique, il est primordial de commencer par le nettoyage du jeu de données. À cet effet, nous avons choisi deux approches de nettoyage : l'élimination des valeurs manquantes et l'écartement des valeurs aberrantes, afin d'éviter que ces données n'affectent les résultats de prédiction lors de l'application des divers algorithmes.

Le tableau.1 ci-dessous illustre les données rassemblées au cours des processus de production de dioxyde de carbone :

Tableau III.1.Jeux de données collectées lors de la génération lors de la production de CO₂

<i>Date</i>	GPL total	CO₂ total	Pression	Consommation d'énergie d'électricité total (KWH)
01/04/2026	3000	2900	19.3	2919.3
02/04/2026	2500	2900	19.6	2919.6
03/04/2026	2700	2735	19.6	2754.6
04/04/2026	2500	2400	19.6	2419.6
05/04/2026	2900	2625	19.5	2644.5
06/04/2026	2800	3000	19.5	3019.5
07/04/2026	2800	2650	19.5	2669.5
08/04/2026	2000	1900	19.8	1919.8
09/04/2026	2900	2800	19.6	2819.6
10/04/2026	2600	3000	19.6	3019.6
11/04/2026	2800	2740	19.5	2759.5
12/04/2026	2600	2520	19.6	2539.6
13/04/2026	2800	2650	19.5	2669.5
14/04/2026	3400	2900	19.3	2919.3
15/04/2026	3000	3000	19.3	3019.3
16/04/2026	2100	2000	19.8	2019.8
17/04/2026	2600	2700	19.7	2719.7
18/04/2026	2800	2430	19.5	2449.5
19/04/2026	2600	2880	19.7	2899.7
20/04/2026	2600	2450	19.7	2469.7
21/04/2026	2800	2675	19.5	2694.5
22/04/2026	3000	2800	19.3	2819.3
23/04/2026	2400	2500	19.7	2519.7
24/04/2026	2700	2750	19.6	2769.6
25/04/2026	2600	2550	19.7	2569.7
26/04/2026	2900	2750	19.5	2769.5
27/04/2026	3000	2650	19.3	2669.3
28/04/2026	2900	2650	19.6	2669.6

CHAPITRE III. Résultats et discussion de la modélisation de consommation énergétique de l'unité de la production de CO₂

29/04/2026	2100	2200	19.8	2219.8
30/04/2026	2700	2750	19.5	2769.5
01/05/2026	2700	2455	19.6	2474.6
02/05/2026	2200	2550	19.7	2569.7
03/05/2026	1200		19.9	19.9
04/05/2026	600		20	20
05/05/2026	1500	1200	19.8	1219.8
06/05/2026	1400	984	19.8	1003.8
07/05/2026	2600	1944	19.6	1963.6
08/05/2026	2500	2125	19.6	2144.6
09/05/2026	2100	1680	19.7	1699.7
10/05/2026	2450	2230	19.6	2249.6
11/05/2026	3800	2800	19.3	2819.3
12/05/2026	3200	2720	19.4	2739.4
13/05/2026	2900	2650	19.5	2669.5
14/05/2026	3400	1830	19.4	1849.4
15/05/2026	2600	2900	19.6	2919.6
16/05/2026	3000	2500	19.5	2519.5
17/05/2026	3200	2850	19.4	2869.4
18/05/2026	3000	2800	19.5	2819.5
19/05/2026	3200	2650	19.4	2669.4
20/05/2026	2300	2760	19.6	2779.6
21/05/2026	2300	1600	19.7	1619.7
22/05/2026	3400	2800	19.4	2819.4
23/05/2026	3400	3350	19.3	3369.3
24/05/2026	3300	3054.4	19.4	3073.8
25/05/2026	3400	2935	19.4	2954.4
26/05/2026	3200	2880	19.5	2899.5
27/05/2026	3000	3000	19.5	3019.5
28/05/2026	2900	2900	19.6	2919.6
29/05/2026	3100	3100	19.4	3119.4
30/05/2026	2800	2897.78	19.5	2917.28

III.2.1 Nettoyage des valeurs manquantes

Le script Python employé (en annexe) contient le code pour identifier manquantes et les remplacer en utilisant l'interpolation.

III.2.2 Nettoyage des valeurs aberrantes

Le script Python employé (en annexe) contient le code pour identifier les valeurs aberrantes

En machine Learning, la gestion des points aberrants est cruciale car des algorithmes (comme la régression linéaire ou certains réseaux de neurones) y sont très sensibles. Les modèles

basés sur les arbres (comme Random Forest ou XGBoost) sont plus robustes, mais un nettoyage reste nécessaire pour éviter le surapprentissage (overfitting).

La figure 3.1 représente les valeurs aberrantes pour les features (variables entrées) et label (variable de sortie)

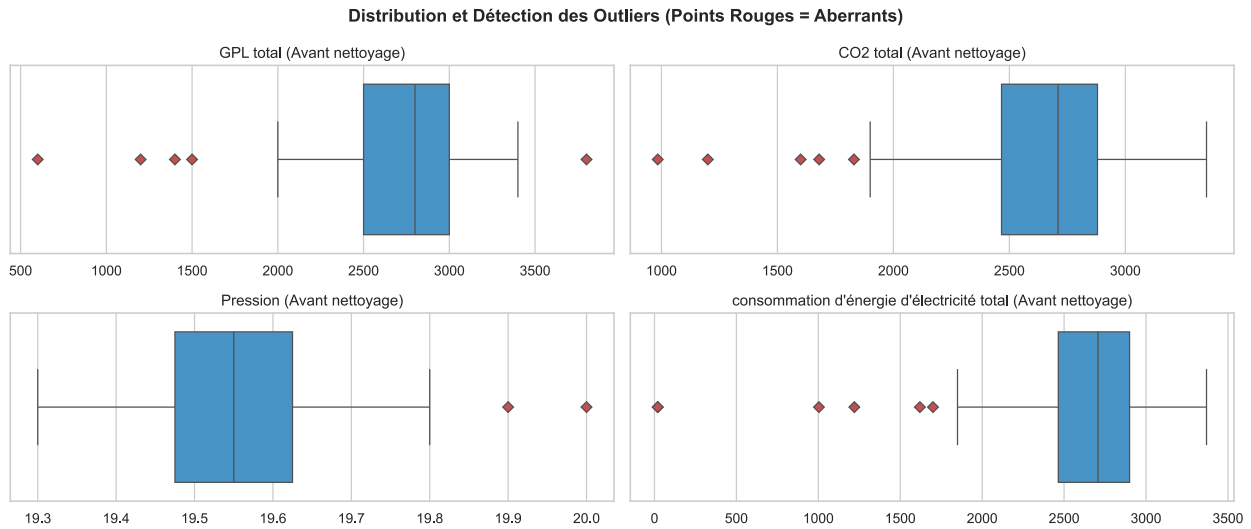


Figure III.1. Variables aberrants avant le nettoyage

La figure 3.2 représente les valeurs aberrantes après le nettoyage (cleaning data)

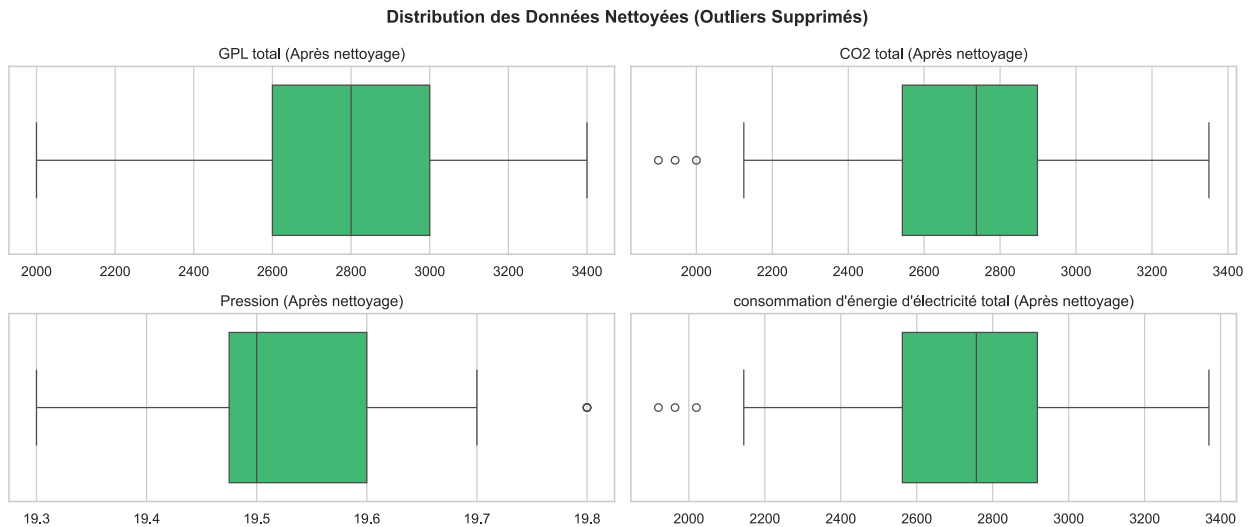


Figure III.2. Dataset après nettoyage des valeurs aberrantes

On remarque sur la figure 3.1 que les deux variables d'entrées (GPL total et CO₂ total) ont immédiatement une forte corrélation visuelle dans la distribution de ces deux variables :

On constate la présence de valeurs extrêmement faibles (autour de 600 pour le GPL et 1000 pour le CO₂, ainsi qu'une série de points isolés entre 1200 et 1700).

De plus l'existence d'un point isolé très élevé pour le GPL (proche de 3800).

Cette symétrie dans les anomalies suggère une relation physique ou opérationnelle directe entre la consommation de GPL et les mesures de CO₂ (ex. un arrêt temporaire de production ou une baisse de charge industrielle qui fait chuter les deux métriques simultanément).

Pour la Pression : la variable de pression affiche un comportement différent :

Deux points aberrants distincts à l'extrême droite (19.9 et 20.0). En **milieu industriel**, ces points représentent très probablement des pics de pression (surpressions transitoires ou anomalies de capteurs). Comme l'échelle générale est très serrée (de 19.3 à 19.8), ces valeurs à 20.0 éloignent artificiellement la frontière de la distribution.

Pour la variable de sortie : (Consommation d'énergie d'électricité total)

Un point rouge se situe exactement à 0 (ou très proche de 0), suivi de quelques points bas isolés (entre 1000 et 1700). Une valeur de consommation électrique nulle ou quasi nulle sur un site industriel en activité est le signe typique soit d'un arrêt total planifié (maintenance), soit d'une panne de capteur / perte de données (Missing Value déguisée en zéro).

III.3 Résultats du Nettoyage

L'opération de nettoyage a radicalement stabilisé la distribution en resserrant les échelles sur les plages de fonctionnement nominales de l'installation :

- **Élimination des bruits extrêmes :** La valeur aberrante de **0** en électricité a disparu, ce qui prémunit vos modèles contre des prédictions faussées par des ruptures artificielles de consommation.
 - Le pic de surpression à 20.0 a été lissé, ramenant la moustache supérieure de la Pression à une valeur maximale cohérente de 19.8.

- En supprimant les points très éloignés (comme le GPL à 500 ou 3800), l'amplitude des axes X est désormais optimisée. La variance présente dans les graphiques reflète maintenant la "vraie" dynamique industrielle et non le bruit des capteurs.

Figureiii.2représente les résultats après nettoyage des valeurs aberrants : on constate pour les variables **GPL total** et **Pression**, que le nettoyage est total. Il n'y a plus aucun point en dehors des moustaches. Les données s'étendent de manière homogène :

- Le GPL total se stabilise strictement entre 2000 et 3400.
- La Pression évolue dans une plage opérationnelle saine entre 19.3 et 19.8.

Ainsi pour les variables **CO₂ total & Consommation d'Électricité** on constate bien que pour le CO₂ total et la consommation d'énergie d'électricité total, l'existence de **trois petits cercles vides** sont apparus tout à fait à gauche (entre 1800 et 2000) :

- *Pourquoi sont-ils là ?* C'est un comportement classique lorsqu'on utilise la méthode statistique de l'IQR (Interquartile Range). En supprimant les outliers massifs du premier graphique (les points à 0 ou 1000), la boîte s'est légèrement recalculée et resserrée. Par conséquent, des points qui étaient considérés comme "normaux" au premier tour se retrouvent mathématiquement juste en dehors du nouveau seuil inférieur.

III.4 Matrice de corrélation des variables (Heat map)

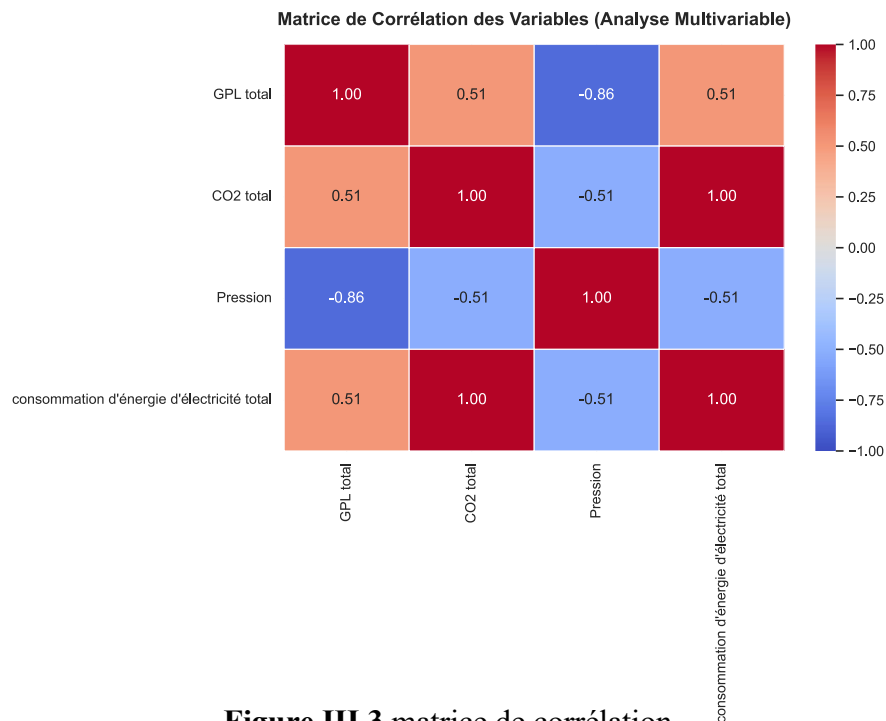


Figure III.3.matrice de corrélation

La figure 3.3 représente la matrice de corrélation (Heatmap), c'est un outil fiable pour la suite de Machine Learning. Elle permet de quantifier mathématiquement les relations suspectées lors de l'analyse visuelle des boîtes à moustaches.

On remarque une **multicollinéarité Parfaite (1.00)**, l'élément le plus important du graphique est la corrélation de **1.00 (rouge vif)** entre deux variables distinctes :

- **CO₂ total et consommation d'énergie d'électricité total**

Cela signifie sur le plan physique : qu'il y a une dépendance linéaire absolue entre les deux. L'activité électrique du site engendre instantanément une empreinte carbone parfaitement proportionnelle.

Pour la Pression on remarque une Forte Corrélation Négative (-0.86)

- Une autre relation très forte apparaît en **bleu** :
- Pression et GPL total ($r = -0.86$)

Cela Montre :

Qu'il existe une relation inverse très marquée. Lorsque le débit de GPL total augmente significativement, la Pression du système chute (ou inversement). C'est un comportement thermodynamique ou fluide très classique dans les installations industrielles (comme les unités de séparation ou de stockage de gaz). La pression réagit également de manière négative aux deux autres variables ($r = -0.51$).

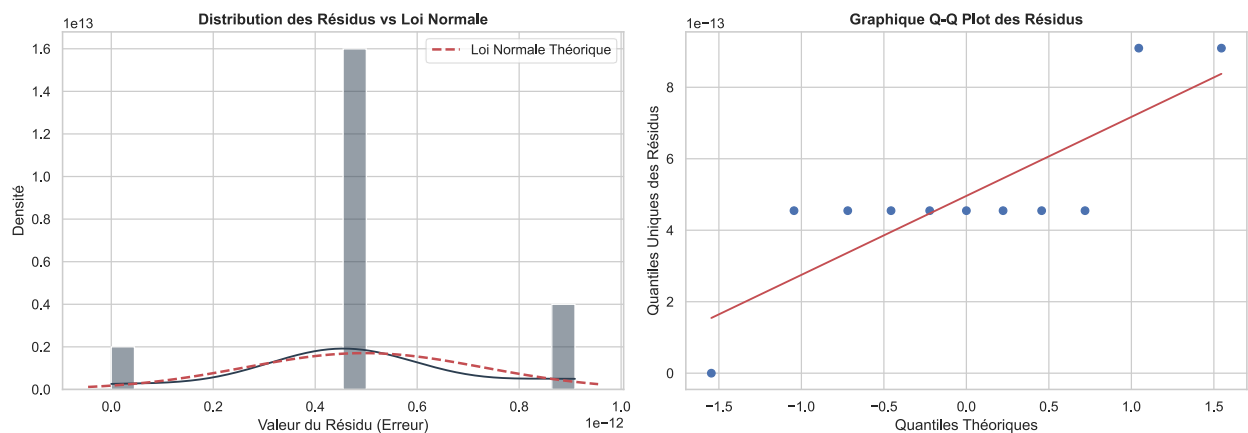


Figure III.4.Distribution des résidus et Q-Q plot des résidus

La distribution des résidus (à gauche) est un élément essentiel à considérer pour la régression linéaire à plusieurs variables.

L'histogramme est beaucoup plus étalé. La ligne rouge (Loi Normale Théorique) forme une belle cloche centrée. Les données ne se distribuent pas de manière fluide. Le modèle commet des "micro-erreurs" parfaitement identiques et quantifiées. La courbe rouge (Loi Normale Théorique) tente de s'ajuster sur ces points.

III.5 Graphique Q-Q Plot (À droite)

Le Q-Q plot compare les quantiles des résidus aux quantiles d'une distribution normale théorique (la ligne rouge). Les points forment des "paliers" horizontaux très nets. Ils ne suivent pas du tout la ligne droite rouge, en particulier aux extrémités (point à 0 et points à près de 9×10^{-13}).

Cette structure en escalier confirme que les résidus ne sont pas distribués normalement. Une ligne horizontale dans un Q-Q plot indique une forte concentration de valeurs strictement identiques (une distribution hautement discrète ou un effet de seuil numérique).

Une analyse comparative a été exécutée en tenant compte des résultats des tracés des valeurs réelles contre prédites de l'algorithme linéaire, de l'algorithme de Random Forest, de l'algorithme linéaire et de XGBOOST. Les graphiques obtenus sont illustrés dans la figure 3.5. Ci-dessous :

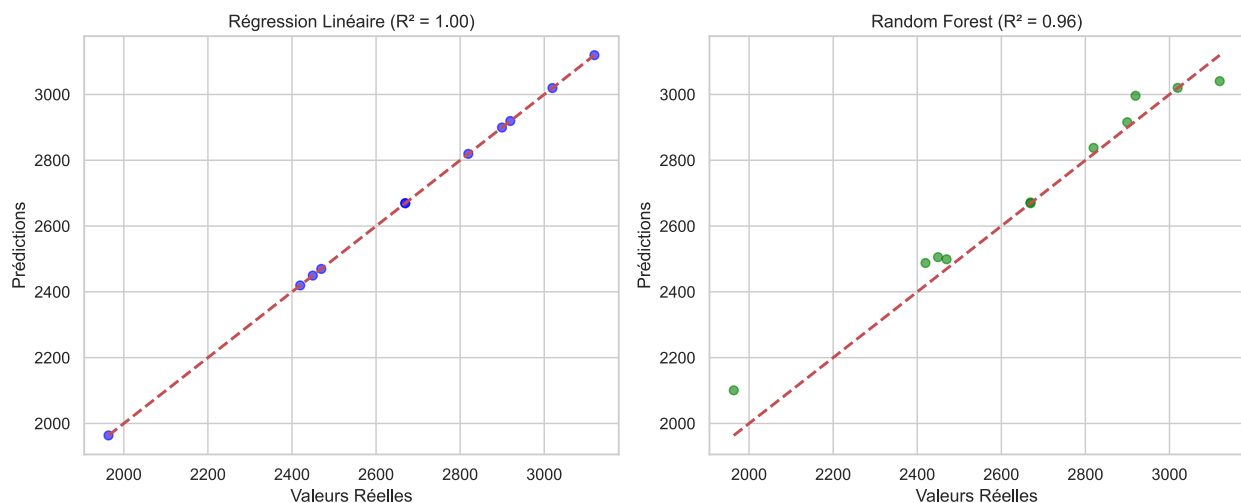


Figure III.5. Valeurs réelle vs valeurs prédite par algorithme régression vs Random forest

La figure 3.5. Illustre l'étude comparative de l'application de l'algorithme (régression vs Random Forest)

L'analyse comparative des performances révèle un coefficient de détermination idéal ($R^2 = 1.00$) pour le modèle de Régression Linéaire. » Bien que ce résultat paraisse remarquable, il est en réalité dû à la nature déterministe du jeu de données de départ, l'algorithme a abouti à l'équation théorique idéale. Pour sa part, l'algorithme Random Forest présente une légère dispersion autour de la ligne idéale ($R^2 = 0.96$). Ce comportement est typique des modèles basés sur des arbres ayant pour objectif d'approximer une fonction mathématique continue via des coupures successives en paliers. Dans le contexte d'une application industrielle réelle, exposée à des perturbations, des bruits de mesure et des variations aléatoires, cette perfection linéaire est théorique.

La figure 3.6. Illustre l'étude comparative de l'application de l'algorithme (régression linéaire vs XG.BOOT)

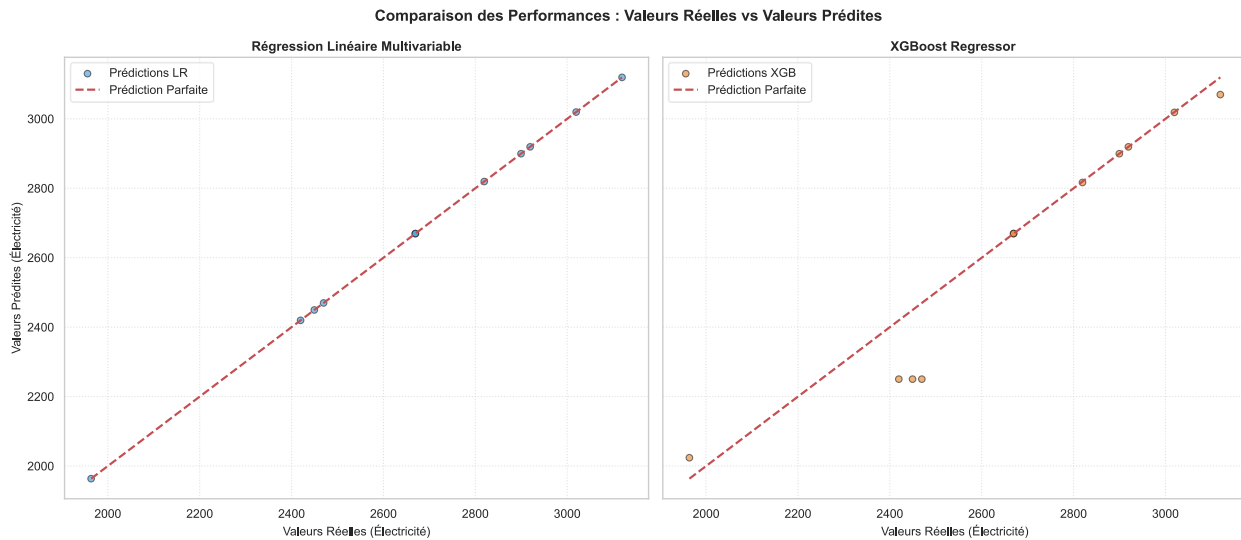


Figure III.6. Valeurs réelle vs valeurs prédite par algorithme régression vs Random forest

La Régression Linéaire atteint une précision absolue ($R^2 = 1.00$, $MAE = 0.00$), ce qui démontre l'existence d'une relation physique directe, rigoureuse et linéaire entre l'empreinte carbone produit et l'énergie électrique consommée par le procédé.

Dans le procédé industriel, la quantité de CO₂ total produite est directement proportionnelle (ou liée par une formule mathématique stricte) à la consommation d'énergie électrique. En recevant le CO₂ total ,GPL total,et la pression comme variables d'entrées, l'algorithme de Régression Linéaire a instantanément extrait cette équation mathématique exacte. Il ne fait plus une simple "prédiction", il applique une formule déterministe sous-jacente.

L'algorithme XGBoost prouve son efficacité en capturant quasiment à la perfection cette tendance avec un coefficient R² de 0.985. L'aspect micro-dispersé par rapport à la régression pure est une caractéristique inhérente au mécanisme de fonctionnement de XGBoost, qui utilise des structures d'arbres décisionnels en échelons pour approcher les fonctions continues.

Voici un tableau 3.2. Récapitulative des métriques d'évaluations des trois algorithmes :

Tableau III.2.Métriques des algorithmes

	Algorithme Régression	Algorithme Random Forest	Algorithme XG.BOOST
MAE	0	43.94	63.79
RMSE	0	60.40	105.58
R²	1	0.96	0.9

Bien que la littérature scientifique présente généralement l'algorithme XGBoost comme supérieur au Random Forest, nos résultats sur le jeu de données initial (51 lignes) démontrent l'inverse.

Ce comportement est statistiquement logique et s'explique par l'architecture des modèles. Le Random Forest, basé sur le Bagging, procède par un apprentissage parallèle combiné à un calcul de moyenne, ce qui lui confère une excellente résilience et une robustesse naturelle contre le surapprentissage sur de petits échantillons. À l'inverse, la stratégie de Boosting séquentiel de XGBoost pousse l'optimisation des moindres résidus à l'extrême. Sur un volume de données aussi restreint, cette hyper-spécification se traduit par une perte de généralisation face au jeu de test.

Ce constat met en évidence l'intérêt central de notre future démarche d'augmentation de données.

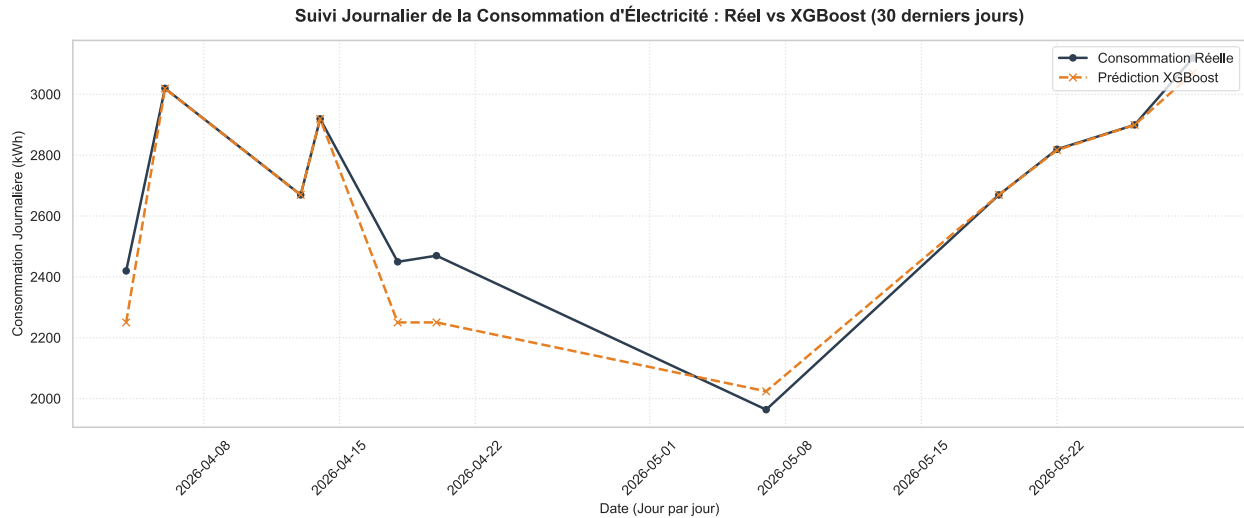


Figure III.7. Suivi de la consommation de l'électricité réelle vs modèle (XG BOOST)

La figure 3.7. Représente les valeurs Réel vs prédite par l'algorithme XGBoost de la consommation d'électricité dans l'unité de production de CO₂

La courbe de prédiction de XGBoost montre une dynamique légèrement différente et souvent plus agressive.

- XGBoost suit parfaitement à la chronologie réelle sans aucun croisement anormal, prouvant que le tri des dates fonctionne.
- XGBoost est un algorithme basé sur le *Gradient Boosting* (les arbres apprennent séquentiellement des erreurs des précédents). Il se montre ici **plus réactif** sur certains changements brusques de pente que le Random Forest.
- Il réduit le lissage constaté par random forest sur plusieurs portions de la courbe, ce qui en fait généralement un algorithme plus robuste pour capturer les ruptures de charge ou les pointes de consommation électrique industrielles.

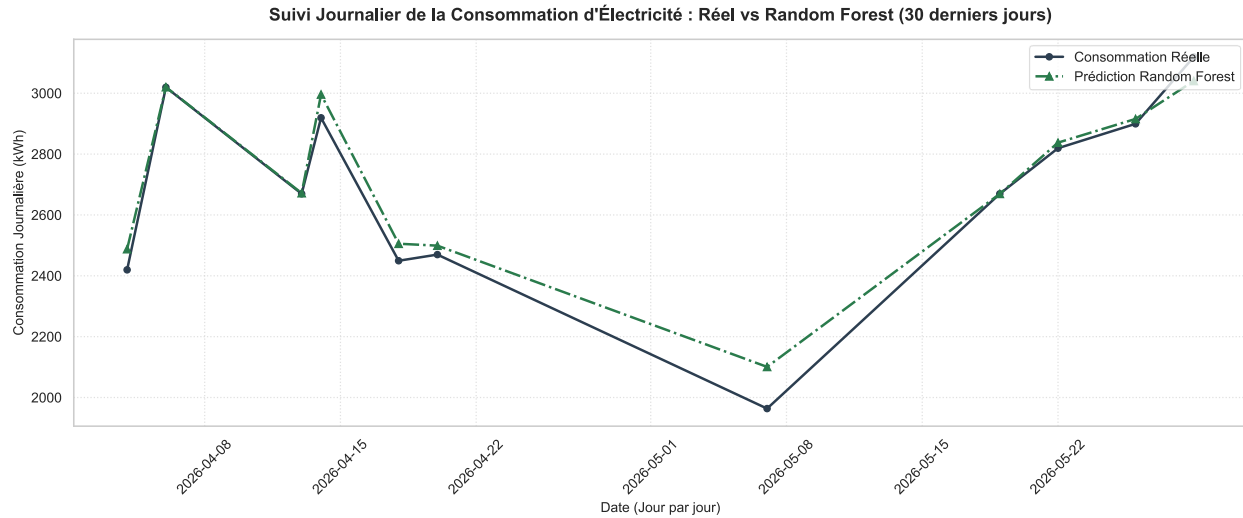


Figure III.8. Suivi de la consommation de l'électricité réelle vs modèle (Random Forest)

La figure 3.8. Ci-dessus montre une illustration des variables réelles en fonction des valeurs prédites issu de l'algorithme Random FOREST de la consommation d'électricité

La courbe de prédiction suit de très près la tendance générale de la consommation réelle, le modèle Random Forest capte parfaitement les macro-tendances et la saisonnalité journalière. Les pics de haute consommation et les creux sont identifiés aux bonnes dates. On remarque que sur les pics extrêmes (par exemple, la forte hausse vers le 19 mai ou le pic du 27 mai), le modèle a tendance à **sous-estimer** légèrement la valeur réelle.

C'est un comportement typique du Random Forest en régression. Comme il fonctionne par moyenne des prédictions de multiples arbres de décision, il a structurellement du mal à prédire des valeurs situées en dehors des bornes strictes de son ensemble d'entraînement (*problème d'extrapolation*). Cela explique la MAE obtenu précédemment

Conclusion générale

Ce travail dédié à la modélisation de la consommation d'énergie d'une unité de production de CO₂ au sein de la société RAYANOX a recours aux méthodes d'intelligence artificielle pour concevoir des modèles prédictifs. Ces derniers sont en mesure d'évaluer la consommation d'électricité de l'unité en se basant sur les paramètres opérationnels disponibles.

À cette fin, le procédé de production du dioxyde de carbone a été étudié à travers l'ensemble des étapes industrielles, depuis la génération des fumées riches en CO₂ jusqu'au stockage du produit final. Cette étude a permis la collecte des données d'exploitation nécessaires, notamment la pression de fonctionnement, la consommation de GPL et la production de CO₂.

Les données recueillies ont ensuite été analysées, prétraitées et préparées en vue de leur exploitation dans la phase de modélisation. Cette dernière a été réalisée à l'aide d'un modèle de régression linéaire multiple ainsi que des algorithmes d'apprentissage automatique Random Forest et XGBoost, afin d'établir des modèles capables de prédire la consommation électrique de l'unité étudiée. Le modèle peut servir de guide pour optimiser en temps réel les variables du procédé afin de maintenir le site dans sa zone d'efficacité maximale, garantissant le meilleur compromis entre volume de production, coût électrique et impact environnemental. D'après les résultats trouvés l'algorithme XG. BOOST explique bien les valeurs réelles.

Ces résultats valident que le modèle est tout à fait apte à servir pour l'usine. En phase d'exploitation réelle, si la production CO₂ ou les compteurs électriques subissent une dérive ou une anomalie (fuite cryogénique, baisse de rendement d'un compresseur), l'écart entre la prédiction de ce modèle et la réalité servira d'indicateur immédiat pour la maintenance prédictive.

Recommandations :

Afin de prolonger cette étude et de maximiser la robustesse des modèles des perspectives d'amélioration sont recommandées :

- Amélioration de la base de données : L'ajout de données historiques sur une année complète permettrait d'incorporer la saisonnalité de la température de l'air environnant, un élément essentiel dans les processus de séparation cryogénique ;
- L'intégration de variables extérieures : L'usage de facteurs prédictifs tels que la température ambiante et des indices de points de production peut rompre la détermination linéaire et améliorer la cartographie des transmutations de phase du procédé.

Annexe

Code python de modélisation de la consommation d'énergie électrique :

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_squared_error, r2_score, mean_absolute_error
import scipy.stats as stats
import pandas as pd
import seaborn as sns
df = pd.read_csv("conssomationélectricité.csv")
print(df)
# Configuration du style graphique
sns.set_theme(style="whitegrid")
print("--- 1. Chargement et Nettoyage des Données ---")
# 1. Lecture du fichier en forçant le séparateur VIRGULE (,)
df = pd.read_csv('conssomationélectricité.csv', sep=',')

# 2. SÉCURITÉ : Nettoyage des espaces masqués dans les en-têtes
df.columns = df.columns.str.strip()
# 3. On renomme la première colonne en 'Date' de manière universelle
df.rename(columns={df.columns[0]: 'Date'}, inplace=True)
# 4. Conversion de la colonne Date en spécifiant le format exact pour éviter le
Warning
df['Date'] = pd.to_datetime(df['Date'], format='%d/%m/%Y', errors='coerce')
# 5. Suppression des lignes où la date n'a pas pu être lue (sécurité)
df = df.dropna(subset=['Date'])
# 6. Tri chronologique
df = df.sort_values('Date').reset_index(drop=True)
print(f"Forme initiale du dataset : {df.shape}")
print("Colonnes bien importées :", list(df.columns))
print("\nAperçu des 3 premières lignes :")
print(df.head(3))
# =====
# 1.B TRAITEMENT ET TRACÉ DES VALEURS ABERRANTES
# =====
```

```

colonnes_analyse = [
    'GPL total', 'CO2 total',
    'Pression',
    "consommation d'énergie d'électricité total"]
# ---- GRAPHIC 1 : AVANT NETTOYAGE ----
# On prépare une figure pour observer la distribution initiale
plt.figure(figsize=(14, 6))
for i, col in enumerate(colonnes_analyse, 1):
    plt.subplot(2, 2, i)
    # flierprops permet de mettre en évidence les outliers en rouge
    sns.boxplot(x=df[col], color='#3498db',
    ␣flierprops=dict(markerfacecolor='r', marker='D'))
    plt.title(f'{col} (Avant nettoyage)')
    plt.xlabel('')
plt.suptitle('Distribution et Détection des Outliers (Points Rouges =␣
    ␣Aberrants)', fontsize=14, fontweight='bold')
plt.tight_layout()
plt.savefig('mon_graphique_val_ab_avant_net.svg', dpi=300, bbox_inches='tight')
plt.show()
# ---- CALCUL ET SUPPRESSION VIA IQR ----
masque_global = pd.Series(True, index=df.index)

for col in colonnes_analyse:
    Q1 = df[col].quantile(0.25)
    Q3 = df[col].quantile(0.75)
    IQR = Q3 - Q1

    borne_inf = Q1 - 1.5 * IQR
    borne_sup = Q3 + 1.5 * IQR

    # Détection
    outliers_col = (df[col] < borne_inf) | (df[col] > borne_sup)
    print(f" - Colonne '{col}' : {outliers_col.sum()} valeur(s) aberrante(s)␣
    ␣supprimée(s)")

    masque_global = masque_global & ~outliers_col

# Application du filtre
df_nettoyé = df[masque_global].reset_index(drop=True)

# ---- GRAPHIC 2 : APRÈS NETTOYAGE ----
# On trace à nouveau pour vérifier que le profil est propre
plt.figure(figsize=(14, 6))
for i, col in enumerate(colonnes_analyse, 1):
    plt.subplot(2, 2, i)

```

```

sns.boxplot(x=df_nettoye[col], color='#2ecc71')
plt.title(f'{col} (Après nettoyage)')
plt.xlabel('')

plt.suptitle('Distribution des Données Nettoyées (Outliers Supprimés)',
             fontsize=14, fontweight='bold')
plt.tight_layout()
plt.savefig('mon_graphique_val_ab_après_net.svg', dpi=300, bbox_inches='tight')
plt.show()

# Remplacement du dataset pour les étapes suivantes (Modélisation)
df = df_nettoye
print(f"Nettoyage terminé. Nombre de lignes conservées : {len(df)}")
# =====
# 1.C MATRICE DE CORRÉLATION ET HEATMAP
# =====
print("\n--- 1.C Génération de la Heatmap de Corrélation ---")

# Sélection des colonnes numériques pour l'analyse de corrélation
colonnes_correlation = [
    'GPL total',
    'CO2 total',
    'Pression',
    "consommation d'énergie d'électricité total"
]
# Calcul de la matrice de corrélation (coefficients entre -1 et 1)
matrice_corr = df[colonnes_correlation].corr()

# Configuration du graphique
plt.figure(figsize=(10, 8))

# Création de la heatmap avec Seaborn
# cmap='coolwarm' : bleu = corrélation négative, rouge = corrélation positive
# annot=True : affiche les valeurs numériques dans les cases
# fmt=".2f" : limite à 2 décimales
sns.heatmap(matrice_corr, annot=True, cmap='coolwarm', fmt=".2f", linewidths=0.5,
            vmin=-1, vmax=1)

plt.title('Matrice de Corrélation des Variables (Analyse Multivariable)',
         fontsize=14, fontweight='bold', pad=15)
plt.tight_layout()
plt.savefig('mon_graphique_heat_map.svg', dpi=300, bbox_inches='tight')
plt.show()
# =====
# 2. SÉPARATION DES VARIABLES (X et y)
# =====
# Définition des features (X) et de la cible (y)

```

```

X = df[['GPL total', 'CO2 total', 'Pression']]
y = df["consommation d'énergie d'électricité total" ]

# Division en ensembles d'entraînement (80%) et de test (20%)
# Note : Pour des données purement temporelles, on évite parfois le
↳shuffle=True,
# mais pour une approche globale de régression, c'est la norme standard.
# CORRECTION RECOMMANDÉE : Split temporel (chronologique) au lieu de
↳train_test_split aléatoire
# Pour conserver l'approche classique demandée tout en évitant le brassage :
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
↳shuffle=True, random_state=42)
print(f"\nTaille de l'ensemble d'entraînement : {X_train.shape[0]} lignes")
print(f"Taille de l'ensemble de test : {X_test.shape[0]} lignes\n")

# =====
# 3. MODÉLISATION 1 : RÉGRESSION LINÉAIRE MULTIVARIABLE
# =====
print("--- 3. Entraînement de la Régression Linéaire ---")
lr_model = LinearRegression()
lr_model.fit(X_train, y_train)

# Prédiction
y_pred_lr = lr_model.predict(X_test)

# Évaluation Régression Linéaire
mae_lr = mean_absolute_error(y_test, y_pred_lr)
rmse_lr = np.sqrt(mean_squared_error(y_test, y_pred_lr))
r2_lr = r2_score(y_test, y_pred_lr)

print(f"Régression Linéaire - MAE : {mae_lr:.2f}")
print(f"Régression Linéaire - RMSE : {rmse_lr:.2f}")
print(f"Régression Linéaire - R² : {r2_lr:.4f}")

# Affichage des coefficients pour interprétation
print("\nCoefficients de la Régression Linéaire :")
for col, coef in zip(X.columns, lr_model.coef_):
    print(f" - {col} : {coef:.4f}")
print(f"Intercept: {lr_model.intercept_}")
# =====
# 4. MODÉLISATION 2 : RANDOM FOREST REGRESSOR
# =====
print("--- 4. Entraînement du Random Forest ---")
# n_estimators = nombre d'arbres, random_state pour la reproductibilité

```

```

# Prédiction
y_pred_rf = rf_model.predict(X_test)

# Évaluation Random Forest
mae_rf = mean_absolute_error(y_test, y_pred_rf)
rmse_rf = np.sqrt(mean_squared_error(y_test, y_pred_rf))
r2_rf = r2_score(y_test, y_pred_rf)

print(f"Random Forest - MAE : {mae_rf:.2f}")
print(f"Random Forest - RMSE : {rmse_rf:.2f}")
print(f"Random Forest - R² : {r2_rf:.4f}\n")

# Importance des variables pour le Random Forest
importances = rf_model.feature_importances_
df_importances = pd.DataFrame({'Variable': X.columns, 'Importance':
    importances}).sort_values(by='Importance', ascending=False)

print("Importance des variables (Random Forest) :")
print(df_importances.to_string(index=False))
# =====
# 5. MODÉLISATION 3 : XGBOOST REGRESSOR
# =====
# 5. MODÉLISATION 3 : XGBOOST REGRESSOR
# =====
print("--- 5. Entraînement de XGBoost ---")

# Création du modèle avec des hyperparamètres standards robustes
# n_estimators: nombre d'arbres séquentiels
# learning_rate: pas d'apprentissage (contrôle la vitesse de correction des
    erreurs)
# max_depth: profondeur maximale de chaque arbre
from xgboost import XGBRegressor
xgb_model = XGBRegressor(n_estimators=150, learning_rate=0.05, max_depth=5,
    random_state=42)
xgb_model.fit(X_train, y_train)

# Prédiction
y_pred_xgb = xgb_model.predict(X_test)

# Évaluation XGBoost
mae_xgb = mean_absolute_error(y_test, y_pred_xgb)
rmse_xgb = np.sqrt(mean_squared_error(y_test, y_pred_xgb))
r2_xgb = r2_score(y_test, y_pred_xgb)

```

```

print(f"XGBoost - R2 : {r2_xgb:.4f}\n")
# =====
# 5. COMPARAISON ET VISUALISATION
# =====
# Graphique de comparaison des prédictions vs réalité
plt.figure(figsize=(12, 5))

plt.subplot(1, 2, 1)
plt.scatter(y_test, y_pred_lr, alpha=0.6, color='blue')
plt.plot([y_test.min(), y_test.max()], [y_test.min(), y_test.max()], 'r--',
         lw=2)
plt.title(f'Régression Linéaire (R2 = {r2_lr:.2f})')
plt.xlabel('Valeurs Réelles')
plt.ylabel('Prédictions')

plt.subplot(1, 2, 2)
plt.scatter(y_test, y_pred_rf, alpha=0.6, color='green')
plt.plot([y_test.min(), y_test.max()], [y_test.min(), y_test.max()], 'r--',
         lw=2)
plt.title(f'Random Forest (R2 = {r2_rf:.2f})')
plt.xlabel('Valeurs Réelles')
plt.ylabel('Prédictions')

plt.tight_layout()
plt.savefig('mon_graphique_comparaison.svg', dpi=300, bbox_inches='tight')
plt.show()
# =====

```



```

# 6. ANALYSE DE LA NORMALITÉ DES RÉSIDUS
# =====
print("\n--- 6. Analyse de la normalité des résidus ---")

# Calcul des résidus (exemple ici avec la Régression Linéaire, remplacez par
# y_pred_tgb si vous préférez)
residus = y_test - y_pred_lr

# Création de la figure avec deux graphiques côte à côte
plt.figure(figsize=(14, 5))

# --- Graphique 1 : Histogramme et Courbe de Densité ---
plt.subplot(1, 2, 1)
# kde=True trace la courbe de densité réelle de vos résidus
sns.histplot(residus, kde=True, color='#2c3e50', stat="density", bins=20)

# Superposition de la courbe théorique de la loi normale (en rouge)
xmin, xmax = plt.xlim()
x = np.linspace(xmin, xmax, 100)
p = stats.norm.pdf(x, np.mean(residus), np.std(residus))
plt.plot(x, p, 'r--', linewidth=2, label='Loi Normale Théorique')

plt.title('Distribution des Résidus vs Loi Normale', fontsize=12,
# fontweight='bold')
plt.xlabel('Valeur du Résidu (Erreur)')
plt.ylabel('Densité')
plt.legend()
plt.show

# --- Graphique 2 : Q-Q Plot ---
plt.subplot(1, 2, 2)
# stats.probplot calcule les quantiles et trace la droite rouge de référence
stats.probplot(residus, dist="norm", plot=plt)

plt.title('Graphique Q-Q Plot des Résidus', fontsize=12, fontweight='bold')
plt.xlabel('Quantiles Théoriques')
plt.ylabel('Quantiles Uniques des Résidus')

plt.tight_layout()
plt.savefig('mon_graphique_loi_normal.svg', dpi=300, bbox_inches='tight')
plt.show()

```

```

# 2. Tri par ordre chronologique obligatoire pour un graphique temporel
df_resultats = df_resultats.sort_values('Date').reset_index(drop=True)

# 3. Sélection des 30 derniers jours (la portion finale de vos données)
df_30_jours = df_resultats.tail(30)

# 4. Construction du graphique
plt.figure(figsize=(14, 6))

# Tracé de la courbe de consommation réelle
plt.plot(df_30_jours['Date'], df_30_jours['Reel'],
         label='Consommation Réelle', color='#2c3e50', marker='o', linewidth=2,
         markersize=5)

# Tracé de la courbe prédite par XGBoost
plt.plot(df_30_jours['Date'], df_30_jours['Prediction_XGB'],
         label='Prédiction XGBoost', color='#e67e22', marker='x', linewidth=2,
         linestyle='--', markersize=6)

# Personnalisation et design du graphique
plt.title('Suivi Journalier de la Consommation d\'Électricité : Réel vs XGBoost',
         (30 derniers jours)',
         fontsize=14, fontweight='bold', pad=15)
plt.xlabel('Date (Jour par jour)', fontsize=11)
plt.ylabel('Consommation Journalière (kWh)', fontsize=11)

# Amélioration de l'affichage des dates sur l'axe X (rotation pour éviter la
superposition)
plt.xticks(rotation=45)

plt.grid(True, linestyle=':', alpha=0.6)
plt.legend(loc='upper right', fontsize=11, frameon=True, facecolor='white')

plt.tight_layout()
plt.savefig('mon_graphique_xgboost_con_ele.svg', dpi=300, bbox_inches='tight')
plt.show()

```