

FACULTE DE TECHNOLOGIE  
DEPARTEMENT  
N° : .....



DOMAINE : ELECTRONIQUE  
D'ELECTRONIQUEFILIERE : ELECTRONIQUE  
OPTION : INSTRUMENTATION

## Mémoire Présenté pour l'obtention du diplôme de Master Académique

Par :

- Rahali Wail Habib Errahmen.
- Mokhtari Akram.
- Oucif Bilal.

Intitulé

### Conception d'une l'Interface MATLAB pour l'Installation des Sites d'Énergies Renouvelables.

Soutenu devant le jury composé de :

Pr. Oudira Houcine.	Université Mohamed Boudiaf - M'SILA	President
Mr. Mezaache Hatem.	Université Mohamed Boudiaf - M'SILA	Encadreur
Dr. Zemouri Nahed.	Université Mohamed Boudiaf - M'SILA	Sous-Encadreur
Dr. Djrioui Mohamed.	Université Mohamed Boudiaf - M'SILA	Examineur

Année universitaire : 2024/2025

# *Dédicaces*

*Nous dédions ce modeste travail :*  
*À nos chers parents qui ont été toujours présents,*  
*merci de fond du cœur pour leurs soutiens, leurs*  
*disponibilités et leurs affections tout au long de nos*  
*études.*  
*À nos chers frères.*  
*À tous nos amis(es) et nos collègues de la promotion*  
*Electronique 2024 - 2025.*

***Wail, Akram et Bilal***

# *Remerciements*

*Avant toute chose, nous remercions Dieu et Le louons pour Son aide et Son soutien tout au long de notre parcours universitaire et pour l'accomplissement de ce mémoire.*

*Nous remercions tout particulièrement **Monsieur Hatem Mezaache et Madame Dr Nahed Zemouri**, Nos encadreurs de mémoire, pour leurs disponibilités, leurs conseils avisés et leurs accompagnements tout au long de ce travail. Leurs expertises et leurs encouragements nous avons permis d'avancer sereinement dans nos recherches.*

*Nous tenons à exprimer notre profonde gratitude aux membres du jury pour l'honneur qu'ils nous ont fait en acceptant d'évaluer ce travail.*

*À tous, un grand merci.*

***Wail, Akram et Bilal***

## **Résumé**

*Notre startup est de développer une interface intelligente sous MATLAB destinée à optimiser l'installation des sites d'énergies renouvelables(solaire). Grâce à l'intelligence artificielle, notre solution analyse les données climatiques, géographiques et techniques afin d'identifier les emplacements optimaux et d'optimiser la configuration des équipements.*

*L'objectif est de réduire les coûts, maximiser la production d'énergie et faciliter la prise de décision pour les entreprises et les collectivités souhaitant investir dans les énergies renouvelables. Cette approche innovante combine algorithmes d'optimisation, apprentissage automatique et modélisation avancée pour une planification précise et efficace.*

*Mots clés : Startup, une interface intelligente, MATLAB, sites d'énergies renouvelables, l'intelligence artificielle, algorithmes d'optimisation, apprentissage automatique.*

## Summary

*Our startup is developing an intelligent interface using MATLAB aimed at optimizing the installation of renewable energy sites (solar). Leveraging artificial intelligence, our solution analyzes climatic, geographic, and technical data to identify optimal locations and optimize equipment configuration.*

*The objective is to reduce costs, maximize energy production, and facilitate decision-making for businesses and communities looking to invest in renewable energy. This innovative approach combines optimization algorithms, machine learning, and advanced modeling for precise and efficient planning.*

**Keywords:** *startup, intelligent interface, MATLAB, renewable energy sites, artificial intelligence, optimization algorithms, machine learning.*

## المخلص

تعمل شركتنا الناشئة على تطوير واجهة ذكية في ماتلاب. تهدف إلى تحسين عملية تركيب مواقع الطاقة المتجددة (الطاقة الشمسية). وبفضل الذكاء الاصطناعي، تقوم حلولنا بتحليل البيانات المناخية والجغرافية والفنية من أجل تحديد المواقع المثلى وتحسين إعدادات المعدات.

الهدف هو تقليل التكاليف، وزيادة إنتاج الطاقة، وتسهيل عملية اتخاذ القرار بالنسبة للشركات والهيئات التي ترغب في الاستثمار في الطاقات المتجددة. وتجمع هذه المقاربة المبتكرة بين خوارزميات التحسين، والتعلم الآلي، والنمذجة المتقدمة لتحقيق تخطيط دقيق وفعال.

الكلمات المفتاحية: شركة ناشئة، واجهة ذكية، مواقع الطاقة المتجددة، الذكاء الاصطناعي، خوارزميات التحسين، التعلم الآلي، ماتلاب.

# *Liste des figures*

## **Chapitre 01 :**

Figure1.1 : Les trois composantes du rayonnement solaire global sur un plan incliné.

Figure1.2 : la composition d'un panneau solaire.

Figure1.3 : Digramme : Le fonctionnement d'un panneau solaire photovoltaïque.

Figure1.4 : Un schéma du fonctionnement de l'énergie thermique.

Figure1.5 : le fonctionnement d'un panneau solaire aérovoltaïque.

Figure1.6 : Diagramme : fonctionnement du l'énergie solaire thermodynamique.

## **Chapitre 02 :**

Figure2.1 : Hiérarchie des domaines de l'Intelligence Artificielle.

Figure 2.2 : Architecture typique d'un perceptron multicouche.

Figure 2.3 : Architecture générale d'un réseau RBF.

Figure2.4 : Principe de fonctionnement d'une Machine à Vecteurs de Support (SVM).

Figure2.5 : Principe de la Régression à Vecteurs de Support (SVR).

Figure2.6 : Architecture typique d'un LSTM.

Figure2.7 : Architecture typique d'un BiLSTM.

## **Chapitre 03 :**

Figure3.1 : Les caractéristiques de PC.

Figure 3.2: Localisation des wilayas de M'sila et Adrar sur la carte géographique de l'Algérie.

Figure 3.3 : les données météorologiques de site Adrar.

Figure 3.4: les données météorologiques de site M'sila.

Figure 3.5 : La corrélation entre l'énergie solaire et les entrées météorologiques de site M'sila.

Figure 3.6 : La corrélation entre l'énergie solaire et les entrées météorologiques de site Adrar.

Figure 3.7 : Comparaison entre la puissance réelle et celle prédite par le modèle SVM- Site Adrar-

Figure 3.8.: Comparaison entre la puissance réelle et celle prédite par le modèle LSTM. - Site Adrar-

Figure 3.9 : Comparaison entre la puissance réelle et celle prédite par le modèle BiLSTM - Site Adrar-

Figure 3.10 : Comparaison entre la puissance réelle et celle prédite par le modèle MLP - Site Adrar-

Figure 3.11 : Comparaison entre la puissance réelle et celle prédite par le modèle SVM- Site M'sila-

Figure 3.12 : Comparaison entre la puissance réelle et celle prédite par le modèle LSTM - Site Msila-

Figure 3.13 : Comparaison entre la puissance réelle et celle prédite par le modèle BiLSTM - Site Msila-

Figure 3.14 : Comparaison entre la puissance réelle et celle prédite par le modèle MLP - Site M'sila-

Figure 3.15 : Interface complète de l'outil de prédiction de la puissance solaire développée sous MATLAB.

Figure 3.16 : Courbes temporelles de la puissance solaire et les variables météorologiques sélectionnées.

Figure 3.17 : Visualisation matrice de corrélation de Pearson entre les différentes variables météorologiques et la puissance solaire.

Figure 3.18 Visualisation de la courbe réelle et prédite de la puissance solaire et affichage des valeurs numériques des métriques de performance.

## *Liste des tableaux*

### ***Chapitre 02 :***

Tableau 2.1: Hyper paramètres de MLP

Tableau 2.2 : Comparaison entre LSTM et BILSTM

### ***Chapitre 03 :***

Tableau3.1 : les paramètres des deux sites.

Tableau3.2 : La corrélation entre l'énergie solaire et les entres météorologiques de site M'sila.

Tableau3.3 : La corrélation entre l'énergie solaire et les entres météorologiques de site Adrar.

Tableau3.4 : les résultats des Critères d'évaluation de site ADRAR.

Tableau3.5 : les résultats des Critères d'évaluation de site MSILA.

# Table des matières

Dédicaces.....	I
Remerciements .....	II
Résumé .....	III
Summary .....	IV
الملخص .....	V
Liste des figures .....	VI
Liste des tableaux .....	VII
Table des matières.....	8
Introduction générale : .....	11
<b>Chapitre 01:Contexte et Fondamentaux des Énergies Renouvelables.....</b>	<b>13</b>
1.Introduction .....	4
2.Les types de Les énergies renouvelables .....	4
2.1.L'énergie éolienne.....	44
2.2. L'énergie hydraulique.....	4
2.3. L'énergie Biomasse .....	4
2.4. L'énergie géothermique.....	5
2.5. L'énergie des déchets.....	5
2.6. Énergie Solaire .....	5
2.6.1. Définition .....	5
2.6.2. Le Soleil.....	6
2.6.3 les Rayons solaires.....	6
2.6.4. Les panneaux solaires .....	8
2.6.4.1.la composition d'un panneau solaire.....	18
2.6.4.2. Le fonctionnement des panneaux solaires.....	19
2.6.5.1. L'énergie solaire photovoltaïque.....	20
2.6.5.1.1. Le fonctionnement d'un panneau solaire photovoltaïque.....	20
2.6.5.1.2. Modules photovoltaïques.....	21
2.6.5.2. L'énergie solaire thermique.....	21
2.6.5.2.1. Les types et le fonctionnement d'un panneau solaire thermique.....	22
2.6.5.2.2. Intérêt du solaire thermique.....	22
2.6.5.3. L'énergie solaire aérovoltaïque.....	22
2.6.5.4. L'énergie solaire thermodynamique.....	23
2.6.5.3.1. Le fonctionnement d'un panneau solaire aérovoltaïque.....	23
2.6.5.4.1. Fonctionnement du l'énergie solaire thermodynamique.....	23
2.6.6. Les principaux avantages de l'énergie solaire.....	14

2.6.7. L'avenir de l'énergie solaire.....	14
3. Conclusion.....	14
<b>Chapitre 02 : L'Intelligence Artificielle appliquée à l'Optimisation des Systèmes d'Énergies Renouvelables. ....</b>	<b>25</b>
1. Introduction.....	16
2. Techniques de prédiction.....	16
2.1. Définition.....	16
2.2. Types de prédiction.....	17
2.2.1. Prédiction numérique (régression).....	17
2.2.2. Prédiction catégorielle (classification).....	17
2.3. Avantages de la prédiction dans le domaine des énergies.....	18
2.4. Séries temporelles et leur rôle dans les systèmes énergétiques.....	18
3. Apprentissage automatique (Machine Learning).....	29
3.1. Les types d'apprentissage.....	29
3.1.1. Apprentissage supervisé.....	29
3.1.2. Apprentissage non supervisé.....	29
3.1.3. Apprentissage semi-supervisé.....	30
3.1.4. Apprentissage auto-supervisé.....	30
4. Intelligence artificielle appliquée à l'énergie renouvelable.....	31
4.1. Définition générale de l'IA.....	32
4.2.1. Réseaux de neurones artificiels.....	23
4.2.1.1. Perceptron multicouche (MLP).....	33
4.2.1.2. Réseau à fonctions de base radiale (RBF).....	35
4.2.1.2.1. Définition.....	35
4.2.2. Apprentissage Automatique (Machine Learning).....	26
4.2.2.1. Machine à vecteurs de support (SVM).....	37
4.2.3. Apprentissage Profond (Deep Learning).....	28
4.2.3.1. Mémoire à long terme (LSTM).....	39
4.2.3.2. Mémoire à long terme bidirectionnelle (BiLSTM).....	40
4.2.3.3. Comparaison entre LSTM et BiLSTM.....	42
5. Conclusion.....	33
<b>Chapitre 03: Conception de l'Interface MATLAB pour l'Installation des Sites d'Énergies Renouvelables - Etude de Cas.....</b>	<b>Error! Bookmark not defined.</b>
1. Introduction.....	<b>Error! Bookmark not defined.</b>
3.2 Description du logiciel MATLAB.....	36
3.3 Collecte et traitement des données météorologiques.....	<b>48</b>
3.3.1 Types de données nécessaires:.....	37
3.3.1.1 Données météorologiques.....	39

3.3.1.1.1 Pression.....	<b>Error! Bookmark not defined.</b>
3.3.1.1.2 Humidité .....	<b>Error! Bookmark not defined.</b>
3.3.1.1.3 Radiation solaire .....	38
3.3.1.1.4 Température .....	38
3.3.1.1.5 Vitesse du vent.....	50
3.3.1.2 Données énergétiques.....	50
3.3.1.2.1 Puissance .....	50
3.3.1.2 Données énergétiques.....	50
3.3.2 Sources des données météorologiques et la puissance réelle .....	50
3.3.2.1 Site d'Adrar.....	50
3.3.2.2 Site de M'sila .....	51
3.3.3 Corrélation entre l'énergie solaire et les entres météorologiques.....	53
3.4 les Critères d'évaluation des modèles de prédiction .....	<b>Error! Bookmark not defined.</b>
3.4.1 L'erreur quadratique moyenne "Root Mean Square Error" (RMSE) : .....	44
3.4.2 l'erreur quadratique moyenne "Mean Squared Error" (MSE): .....	45
3.4.3 L'erreur quadratique moyenne en pourcentage " Mean Absolute Percentage Error" (MAPE):.....	45
3.4.4 l'erreur absolue moyenne " Mean Absolute Error" (MAE):.....	45
3.4.5 coefficient de détermination ( $R^2$ ) : .....	46
3.5 Étude de cas : Application pratique sur un site d'énergie renouvelable.....	56
3.6 Présentation et analyse des résultats expérimentaux.....	57
3.6.1 Site de Adrar .....	57
3.6.1.1 Résultats numériques des modèles de prédiction.....	57
3.6.1.2 Visualisation des courbes de prédiction.....	57
3.6.1.3 Analyse comparative des modèles pour le site d'Adrar.....	60
3.6.2 Site de M'sila.....	61
3.6.2.1 Résultats numériques des modèles de prédiction.....	61
3.6.2.2 Analyse comparative des modèles pour le site de M'sila.....	64
3.7 Conception de l'interface graphique MATLAB pour la prédiction de puissance.....	65
3.7.1 Objectifs de l'interface.....	65
3.7.2 Architecture et composants de l'interface.....	65
3.7.2.1 Bloc pour le chargement des Données Réelles.....	66
3.7.2.2 Bloc de Visualisation.....	67
3.7.2.3 Bloc de Choix du Modèle de Prédiction.....	68
3.7.2.4 Bloc Résultats de la Prédiction.....	68
3.7.4 Fonctionnalités principales.....	69
<b>3.8 Conclusion .....</b>	<b>70</b>
Conclusion générale.....	71
Références :.....	72

## ***Introduction générale***

Dans le monde d'aujourd'hui, l'énergie est nécessaire pour garantir un bon niveau de vie. Il est donc essentiel de disposer de sources d'énergie abordables qui n'entraînent ni pollution ni changement climatique. Les énergies renouvelables sont la solution, il estime qu'il est désormais temps de passer des combustibles fossiles à ce type d'énergies .

Heureusement, une partie de l'énergie que nous utilisons n'a pas ces conséquences néfastes, en particulier l'énergie générée par le solaire, l'éolien et l'hydroélectricité. Dans de nombreuses régions du monde, les générateurs solaires et éoliens sont en passe de devenir la source d'énergie la moins chère, et une alternative viable aux combustibles fossiles. De plus, ces sources d'énergie sont renouvelables, car elles se reconstituent naturellement en quelques jours ou décennies. Lorsque l'énergie produite est abordable et que sa production ne nuit pas à l'environnement ou aux personnes, l'approvisionnement en énergie renouvelable est durable. Cette énergie durable pourrait offrir de nombreux avantages immédiats. Elle pourrait réduire la pollution atmosphérique qui sévit aujourd'hui dans de nombreuses villes, fournir une énergie moins chère et de nombreux nouveaux emplois, et assurer la sécurité énergétique de millions de personnes à un coût abordable [1].

En raison de l'augmentation de la population, il est devenu nécessaire de trouver des solutions viables pour la production d'électricité au lieu de dépendre des sources de combustibles telles que le pétrole, le charbon et le gaz naturel. En effet, ces ressources ne suffisent plus à répondre aux besoins croissants du monde, en plus de leurs effets néfastes sur l'environnement. Dans ce contexte, les énergies renouvelables représentent la meilleure solution pour l'avenir, car elles sont plus durables et moins polluantes. Parmi ces sources, l'énergie éolienne et l'énergie solaire constituent les alternatives les plus importantes en Algérie, en raison de l'abondance de l'ensoleillement tout au long de l'année. Les cellules photovoltaïques et les éoliennes figurent parmi les meilleures technologies pour produire de l'électricité, offrant une solution efficace et économique au problème énergétique. Cela en fait une option prometteuse pour assurer un développement durable dans le pays [2].

Les énergies renouvelables, grâce à leur disponibilité illimitée, représentent une opportunité pour plus de 2,5 milliards de personnes à travers le monde, en particulier dans les pays du tiers-monde et les régions isolées, d'accéder à l'électricité. Ces avantages, associés aux progrès technologiques constants, encouragent davantage d'investissements dans le secteur des énergies renouvelables de la part des grandes puissances industrielles, renforçant ainsi l'avenir d'une énergie propre et durable [3].

**Plan de mémoire :** Ce manuscrit commence par une introduction générale. Par la suite, il est réparti en trois chapitres :

**Le premier chapitre :** Dans ce chapitre, nous présentons brièvement les types d'énergies renouvelables avec un accent particulier sur l'énergie solaire.

**Le deuxième chapitre :** nous présentons dans ce chapitre la définition de l'intelligence Artificielle et ses types. Nous terminerons le chapitre par Les Techniques de Prédiction (MLP, SVM, LSTM, BILSTM).

**Le troisième chapitre :** consacré à l'implémentation et la réalisation de la proposition de notre programme, et aussi les expérimentations, les résultats obtenus et les comparaisons avec des travaux similaires. Le mémoire se termine par une conclusion générale et quelques perspectives.

# **CHAPITRE 01**

## Contexte et Fondamentaux des Énergies Renouvelables

## 1.Introduction

Dans ce chapitre, nous présentons brièvement les types d'énergies renouvelables (Énergie éolienne, hydraulique, Biomasse, géothermique, l'énergie des déchets), Et des Généralités sur les énergies renouvelables. Avec un accent particulier sur l'énergie solaire.

Les énergies renouvelables désignent des énergies issues de sources naturelles qui se régénèrent à un taux plus élevé que leur utilisation. Des sources qui se renouvellent continuellement incluent, par exemple, la lumière solaire et le vent. Les énergies renouvelables sont omniprésentes et foisonnent dans notre environnement. L'énergie renouvelable produit nettement moins d'émissions que l'utilisation de combustibles fossiles. Pour lutter contre la crise climatique, il est essentiel de substituer les combustibles fossiles, qui sont actuellement responsables d'une grande part des émissions, par des sources d'énergies renouvelables. Dans la majorité des pays, les énergies renouvelables sont aujourd'hui moins coûteuses et contribuent à générer trois fois plus d'emplois que les énergies fossiles.[4]

## 2. Les types de Les énergies renouvelables

### 2.1. L'énergie éolienne

L'énergie éolienne est une source d'énergie renouvelable. C'est une ressource énergétique intermittente qui n'est pas générée à la demande, mais plutôt en fonction des conditions climatiques. Par conséquent, elle requiert des infrastructures de production ou de stockage alternatives pendant les moments où elle est indisponible. Celles-ci peuvent être prévues avec une assez bonne précision.

L'énergie éolienne représente l'énergie produite par le vent, dont la puissance (énergie cinétique) est exploitée pour propulser des voiliers et autres moyens de transport ou convertie grâce à un mécanisme aérogénérateur, tel qu'une éolienne ou un moulin à vent, en une énergie pouvant être utilisée de multiples façons.[5]

### 2.2. L'énergie hydraulique

L'énergie hydraulique permet de fabriquer de l'électricité, dans les centrales hydroélectriques, grâce à la force de l'eau. Cette force dépend soit de la hauteur de la chute d'eau (centrales de haute ou moyenne chute), soit du débit des fleuves et des rivières (centrales au fil de l'eau). C'est une énergie qui n'émet pas de gaz à effet de serre, elle est utilisable rapidement grâce aux grandes quantités d'eau stockée et c'est une énergie renouvelable très économique à long terme.[6]

### 2.3. L'énergie Biomasse

D'un point de vue énergétique, la biomasse représente toute masse vivante à partir de laquelle de l'énergie peut être obtenue par combustion ou fermentation[7][8]. L'énergie tirée directement de la biomasse est parfois aussi appelée « bioénergie » ; ce terme exclut ainsi les combustibles fossiles, également issus de la biomasse et transformés sur plusieurs milliers d'années [9].

La biomasse est considérée comme une énergie renouvelable tant que la quantité de matière utilisée est égale ou inférieure à la quantité qui peut être régénérée. La biomasse peut être séparée en deux catégories : traditionnelle et moderne. La biomasse traditionnelle englobe la combustion de bois énergie, d'excréments d'animaux et de charbon de bois, tandis que la biomasse moderne concerne des procédés technologiques tels que la production de granulés de bois ou de biocarburants.

## **2.4. L'énergie géothermique**

La production d'électricité géothermique nécessite le forage de puits, souvent à des profondeurs de 1,6 kilomètre ou davantage, dans les réservoirs souterrains. Ces puits permettent d'extraire de la vapeur et de l'eau hautement chaudes qui sont ensuite utilisées pour faire fonctionner des turbines connectées à des générateurs électriques. La première production d'électricité géothermique a eu lieu à Larderello, en Italie, en 1904.

On distingue trois catégories de centrales géothermiques.: à vapeur sèche, flash et binaire. La vapeur sèche, la plus ancienne technologie géothermique, extrait la vapeur des fractures du sol et l'utilise pour entraîner directement une turbine. Les installations de flashage tirent de l'eau chaude profonde, à haute pression, vers de l'eau plus froide, à basse pression. La vapeur qui résulte de ce processus est utilisée pour entraîner la turbine. Dans les installations binaires, l'eau chaude passe par un fluide secondaire avec un point d'ébullition beaucoup plus bas que l'eau. Cela provoque la transformation du fluide secondaire en vapeur, qui entraîne alors une turbine. La plupart des centrales géothermiques du futur seront des centrales binaires.

La géothermie est générée dans plus de vingt nations. Les États-Unis sont le premier producteur mondial d'énergie géothermique, et le plus grand développement de ce type au monde est celui de The Geysers, qui se situe au nord de San Francisco en Californie. En Islande, une multitude de structures et même de piscines sont réchauffées grâce à l'eau chaude géothermique. On dénombre au moins 25 volcans actifs en Islande, sans oublier de nombreuses sources thermales et geysers [10].

## **2.5. L'énergie des déchets**

L'énergie des déchets, aussi appelée valorisation énergétique des déchets, désigne le processus par lequel les déchets non recyclables sont transformés en énergie (chaleur, électricité ou carburant) grâce à des procédés comme l'incinération, la méthanisation ou la gazéification. Ce procédé permet à la fois de réduire le volume des déchets et de produire de l'énergie utile [11].

## **2.6. Énergie Solaire**

### **2.6.1. Définition**

L'énergie solaire est une forme d'énergie renouvelable qui convertit la lumière du soleil en électricité à l'aide de panneaux photovoltaïques. Elle est de plus en plus utilisée pour sa capacité à réduire les émissions de gaz à effet de serre et sa disponibilité quasi illimitée.

Elle est une source d'énergie qui est dépendante du soleil. Cela signifie que la matière première est le soleil. Elle se place dans la catégorie des énergies renouvelables puisqu'on la considère comme inépuisable. On dit aussi que c'est une énergie 100% verte car sa production n'émet que peu de CO<sub>2</sub>.

Grâce à cette énergie, il est possible de produire de l'électricité. Elle sera captée par des panneaux solaires ou des centrales thermiques. Ces installations captent les rayons produits par le soleil. Elles convertissent ensuite l'énergie du soleil en électricité ou en chaleur. Les panneaux solaires qui captent cette énergie peuvent être installés sur différentes structures : des ombrières solaires, des hangars photovoltaïques industriels ou agricoles, au sol, etc [12].

## 2.6.2. Le Soleil

Le soleil, une ressource d'énergie presque infinie, qui envoie à la surface de la terre un rayonnement qui représente chaque année environ 8400 fois la consommation énergétique de l'humanité. Cela correspond à une puissance instantanée reçue de 1 kw crête par mètre carré (kwc/m<sup>2</sup>) répartie sur tout le spectre, de l'ultraviolet à l'infrarouge. Les déserts de notre planète reçoivent en 6 heures plus d'énergie du soleil que n'en consomme l'humanité en une année.

S. Petibon, "Nouvelles architectures distribuées de gestion et de conversion de l'énergie pour les applications photovoltaïques" [13].

## 2.6.3les Rayons solaires

Le rayonnement solaire désigne l'ensemble des ondes électromagnétiques émises par le Soleil. Il se compose donc d'ultraviolets, de la lumière visible, mais également d'ondes radio en plus de rayons cosmiques. Il est une source d'énergie primaire indispensable à de nombreux processus terrestres. Il fournit l'énergie nécessaire à la photosynthèse, qui est le fondement de la chaîne alimentaire. De plus, l'énergie solaire capturée par les panneaux photovoltaïques permet de produire de l'électricité, réduisant ainsi la dépendance aux combustibles fossiles et contribuant à la lutte contre le réchauffement climatique.

### 2.6.3.1. Les caractéristiques du rayonnement solaire sur terre

La quantité d'énergie solaire reçue à un endroit spécifique de la planète est influencée par :

- L'énergie solaire envoyée par le Soleil, qui fluctue avec son activité de fusion.
- La nébulosité (nuages, brouillards, etc.), qui est par exemple importante à l'équateur et plus faible en milieu intertropical.
- La latitude, la saison et l'heure, qui influent sur la hauteur du soleil dans le ciel et donc l'énergie reçue par unité de surface au sol.

Par conséquent, les éléments qui ont un impact sur la production effective sont :

- **L'enseulement** : La quantité de lumière solaire reçue par le panneau dépend de la localisation géographique, du climat et de la saison.
- **L'orientation et l'inclinaison** : L'angle et la direction des panneaux par rapport au soleil affectent leur efficacité.

- **La température** : Les performances des panneaux photovoltaïques diminuent généralement avec des températures élevées.

- **L'ombrage** : Les ombrages causés par des bâtiments, des arbres ou d'autres obstacles peuvent réduire la production d'électricité.

- **La propreté** : La poussière, les feuilles et autres débris sur les panneaux peuvent également réduire leur efficacité [14].

### 2.6.3.2. Composition du rayonnement solaire

Suite à l'atténuation du rayonnement solaire par les différents phénomènes atmosphériques, le rayonnement solaire sur une surface au sol est composé comme suit :

- **Rayonnement direct** :

C'est la fraction du rayonnement solaire qui arrive directement au sol, dont le trajet est « linéaire » (il y a, en fait, de légères déviations) et unique à un instant donné. Si  $I$  est le rayonnement direct.

- **Rayonnement diffus** :

Le rayonnement diffus est le rayonnement provenant de toute la voûte céleste. Ce rayonnement est dû à l'absorption et à la diffusion d'une partie du rayonnement solaire par l'atmosphère et à sa réflexion par les nuages. Ainsi, par ciel serein, il constitue 20% de l'énergie globale. Par ciel couvert, il correspond à la totalité de l'énergie reçue au sol.

- **Rayonnement réfléchi (L'albédo)** :

Le rayonnement solaire réfléchi est le rayonnement qui est réfléchi par le sol ou par des objets se trouvant à sa surface. Ce rayonnement dépend de l'albédo du sol et il peut être important lorsque le sol est particulièrement réfléchissant (eau, neige).

- **Rayonnement global** :

Une surface horizontale C'est l'ensemble du rayonnement d'origine solaire qui parvient sur une surface horizontale sur le globe terrestre. Il comprend donc la composante verticale du rayonnement solaire direct et rayonnement solaire diffus. Une surface inclinée Rayonnement global sur une surface inclinée est la somme des rayonnements : Direct, Diffus et Réfléchi.

Rayonnement global = Rayonnement direct+ Rayonnement diffus+ Rayonnement réfléchi [15].

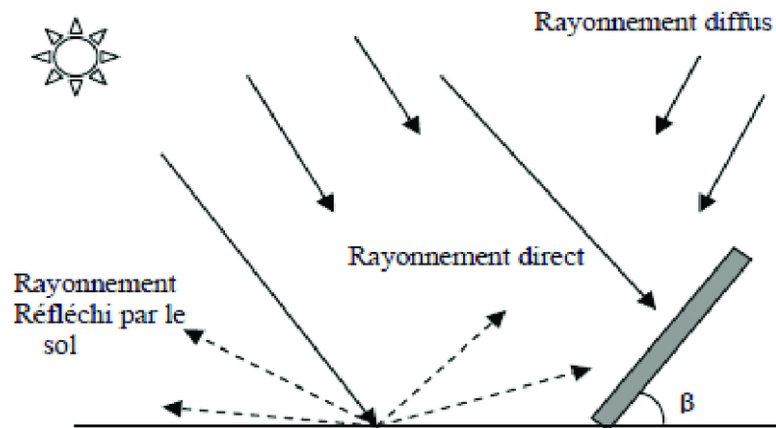


Figure1.1 : Les trois composantes du rayonnement solaire global sur un plan incliné.

## 2.6.4. Les panneaux solaires

Un panneau solaire est cet équipement installé au niveau du sol, sur les toitures ou sur les structures d'ombrage de stationnement. Par exemple qui utilise la lumière du soleil pour produire de l'énergie électrique ou de la chaleur (énergie thermique qui a pour but de réchauffer l'intérieur des bâtiments ou de fournir de l'eau chaude).

C'est cet appareil placé au sol, sur les toits ou sur les ombrières qui utilise la lumière du soleil pour produire de l'énergie électrique ou de la chaleur (énergie thermique qui a pour but de réchauffer l'intérieur des bâtiments ou de fournir de l'eau chaude) [16].

Il existe trois types :

- **Les panneaux solaires thermiques (capteurs solaires thermiques)** : Ils transforment la lumière en chaleur, ensuite utilisée pour produire de l'eau chaude.
- **Les panneaux solaires photovoltaïques (installations photovoltaïques)** : Ils transforment directement la lumière en électricité solaire.
- **Panneaux solaires hybrides (PV-T)** : Combinent photovoltaïque + thermique : produisent de l'électricité et de la chaleur, Son installation est plus complexe.

### 2.6.4.1. la composition d'un panneau solaire

La composition du panneau solaire comprend plusieurs parties :

La partie mécanique est un ensemble de cellules photovoltaïques encapsulées dans de l'EVA, un châssis en aluminium, une coque arrière, une plaque de verre côté soleil. Et la partie électrique permet de connecter le panneau solaire à l'installation.

**a) Les cellules photovoltaïques** : Le silicium parmi les matériaux principaux utilisés dans la fabrication des cellules photovoltaïques. Qu'il s'agisse de panneaux photovoltaïques monocristallins ou polycristallins, les cellules qui en résultent sont composées d'un mélange de silicium ou non. Chaque cellule

solaire subit un processus de dopage afin d'être capable de produire un courant électrique. Les cellules sont connectées en série ou en parallèle, formant ainsi un système de génération d'électricité.

**b) Le dos du panneau solaire :** Petits panneaux solaires, panneau solaire monocristallin, panneau solaire polycristallin, la face arrière doit les protéger des agressions (UV, amplitudes de température, produits chimiques). Pour le dos, les matériaux de composition du panneau solaire sont des polymères. Le Tedlar permet la fabrication de cette partie tournée vers la toiture. L'épaisseur est entre 1 et 2 mm, et la couleur du Tedlar pour panneaux est souvent noire.

**c) La surface du panneau photovoltaïque :** Dans la composition du panneau solaire, le verre trempé est la face côté soleil. Pour les panneaux photovoltaïques, il a des qualités de protection idéales (chocs, UV, humidité). Le nettoyage de panneau solaire suppose que cette partie du panneau soit toujours soignée. Cela permet de bénéficier du rayonnement du soleil maximum, et de garantir le maintien du rendement du panneau photovoltaïque. D'autres matériaux comme le plexiglas et le polycarbonate existent.

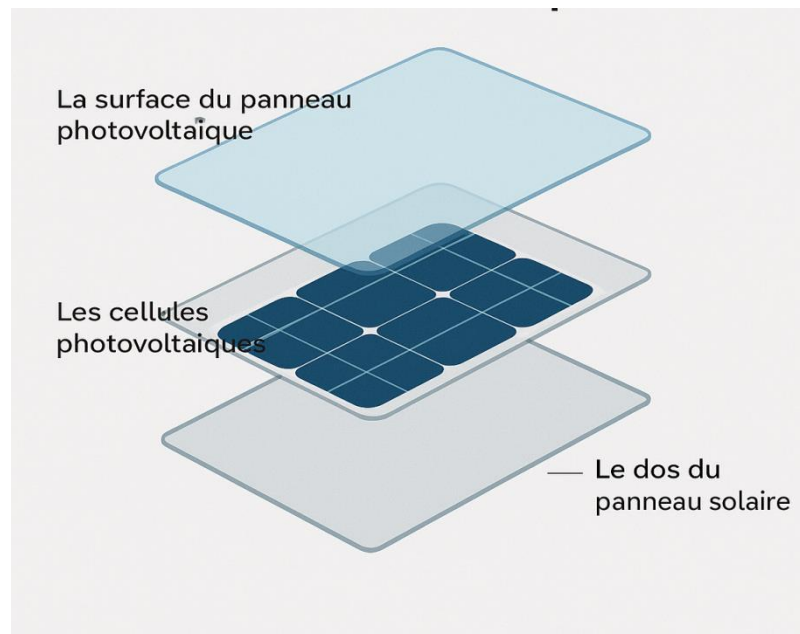


Figure1.2 : la composition d'un panneau solaire.

#### 2.6.4.2. Le fonctionnement des panneaux solaires

Un panneau solaire est capable de convertir une partie du rayonnement solaire en électricité. Lorsque la lumière frappe la surface d'un panneau solaire, les cellules photovoltaïques réagissent en libérant des électrons : un courant électrique continu se crée.

Le courant continu n'est pas utilisable en l'état. C'est pourquoi un onduleur transforme le courant continu en courant alternatif [17].

### 2.6.5.1. L'énergie solaire photovoltaïque

La plus répandue et la plus connue des énergies solaires est l'énergie photovoltaïque. C'est celle qui est employée pour l'autoconsommation. Il s'agit de produire de l'électricité grâce à des panneaux solaires photovoltaïques.

#### 2.6.5.1.1. Le fonctionnement d'un panneau solaire photovoltaïque

Les étapes suivantes sont impliquées dans la production d'énergie photovoltaïque :

- 1) Sous l'action des rayons du soleil, les cellules photovoltaïques produisent un courant continu.
- 2) Le courant continu est transformé par un onduleur solaire en courant alternatif.
- 3) Ce courant alternatif est injecté dans le réseau et alimente en électricité les appareils électriques (box internet, électroménager, ampoules etc.) de votre foyer.
- 4) Si l'auto consommateur le souhaite, le surplus d'électricité peut être vendu au réseau en obligation d'achat (OA Solaire). Enedis d'injecter l'excédent de production sur son réseau et d'approvisionner les bâtiments aux alentours.



Figure 1.3 : Digramme de fonctionnement d'un panneau solaire photovoltaïque.

#### 2.6.5.1.2. Modules photovoltaïques

Il existe plusieurs techniques de modules solaires photovoltaïques :

- Les modules solaires monocristallins possèdent le meilleur rendement au mètre carré et sont essentiellement utilisés lorsque les espaces sont restreints.
- Les modules solaires polycristallins sont moins chers mais présentent un rendement plus faible.
- Les modules solaires amorphes ont aussi des atouts car ils peuvent être souples et ont une meilleure production par faible lumière. Cependant, le silicium amorphe possède un rendement divisé par deux par rapport à celui du cristallin, cette solution nécessite donc une plus grande surface pour la même puissance installée. Toutefois, le prix au mètre carré installé est plus faible que pour des panneaux composés de cellules cristallines [18].

### 2.6.5.2. L'énergie solaire thermique

L'énergie solaire thermique est la chaleur produite par le rayonnement solaire. Elle est recueillie dans l'intention de réchauffer un fluide (qu'il soit liquide ou gazeux). Le fluide peut ensuite exploiter l'énergie reçue de manière directe (chauffage, eau chaude sanitaire, etc.) ou indirectement (création de vapeur d'eau pour alimenter des alternateurs et produire de l'électricité, génération de refroidissement, etc.).

Il s'agit donc d'une forme d'énergie solaire, produite par le Soleil sous forme de rayonnement. Une autre forme importante est l'énergie solaire photovoltaïque, qui exploite l'effet photoélectrique pour convertir les photons émis par le Soleil en électricité.

Le capteur solaire thermique est un dispositif employé pour convertir l'énergie solaire en chaleur. Cette génération d'énergie repose sur des principes physiques essentiels tels que l'absorption et la transmission de chaleur. Concernant spécifiquement les systèmes de concentration (comme la centrale solaire thermodynamique, le four solaire, etc.), le phénomène de réflexion a également une grande importance.

L'énergie solaire thermique connaît une progression, bien que celle-ci soit considérablement ralentie depuis 2014. Avec une part de 72,8 % de la capacité des capteurs en place, la Chine se démarque nettement. Cependant, lorsqu'il s'agit de puissance par habitant, elle est surpassée par sept nations, y compris Chypre, l'Autriche, la Grèce et l'Australie [19].

### 2.6.5.2.1. Les types et le fonctionnement d'un panneau solaire thermique

Un panneau solaire thermique est composé de capteurs qui viennent récupérer la chaleur du soleil. Cela permet de faire monter en température un fluide caloporteur. Ce fluide va alors permettre de chauffer :

- L'eau du circuit de chauffage pour qu'elle arrive jusque dans vos radiateurs.
- L'eau présente dans la cuve d'un ballon d'eau chaude.

Lorsque le soleil vient à manquer un système comme une résistance électrique ou une chaudière d'appoint vient prendre le relais pour assurer la continuité de la production de chaleur.

Si les panneaux thermiques présentent tous le même fonctionnement, ils ne sont pas tous fabriqués de la même manière. Il existe 3 grands types de panneaux solaires thermiques :

Les capteurs plans opaques, Les capteurs plans vitrés, Les capteurs à tubes sous vide.

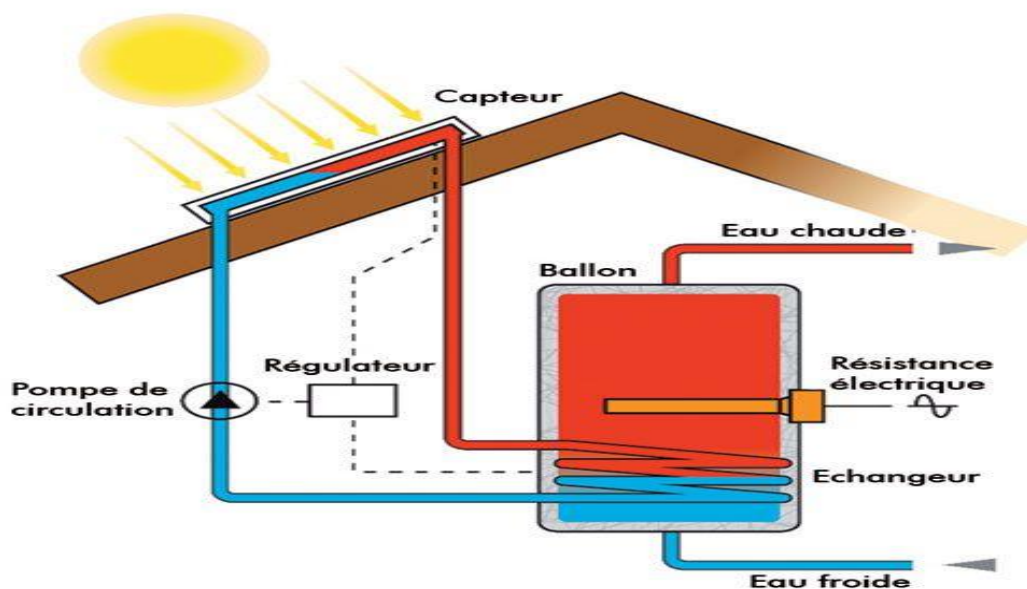


Figure1.4 : Un schéma du fonctionnement de l'énergie thermique.

### 2.6.5.2.2. Intérêt du solaire thermique

Le captage de l'énergie solaire thermique a plusieurs avantages. En ce qui concerne le chauffage, les capteurs destinés au chauffage sont relativement simples, rustiques et durables.

Concernant l'électricité, les systèmes de turbine à vapeur classiquement utilisés reposent sur des composants parfaitement sûrs et éprouvés. Dans les zones très ensoleillées, la rentabilité est prouvée : une centrale solaire thermodynamique est au Maroc amortie au niveau énergétique en cinq mois[20].(c'est-à-dire qu'elle aura produit plus d'énergie qu'en a nécessité sa construction et son démarrage), ce qui est comparable à l'éolien (quatre à sept mois), mais bien plus rapide que les modules photovoltaïques en silicium (qui nécessitent en 2009 trois à cinq ans pour rembourser leur « dette énergétique », mais qui nécessiteront moins de frais et de travail pour leur entretien et fonctionnement). Enfin, un fort potentiel de développement est disponible dans plusieurs pays en développement, qui présentent un impact a priori modéré sur l'environnement (déserts, zones arides, etc.). Selon le DLR, une capacité de plus de 3 GW est réaliste en Europe, tandis qu'elle serait de 15 GW pour toute la planète [21].

Dans le domaine de la purification de l'eau, le distillateur solaire peut fournir de l'eau potable en suffisance dans les zones tropicales et subtropicales.

### 2.6.5.3. L'énergie solaire aérovoltaique

L'énergie solaire aérovoltaique est une technologie hybride qui associe la production d'électricité photovoltaïque à la récupération de la chaleur générée sous les panneaux solaires. Elle permet ainsi une double valorisation de l'énergie solaire : électrique et thermique [22].

#### 2.6.5.3.1. Le fonctionnement d'un panneau solaire aérovoltaique

Les panneaux aérovoltaiques sont des panneaux solaires photovoltaïques, dont la face arrière est ventilée afin de récupérer puis de diffuser de l'air chaud dans votre habitation.

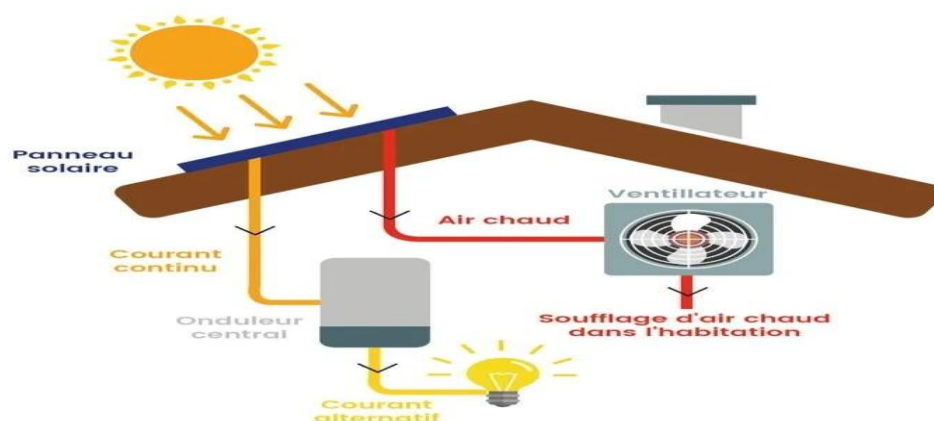


Figure 1.5 : le fonctionnement d'un panneau solaire aérovoltaique.[23]

### 2.6.5.4. L'énergie solaire thermodynamique

L'énergie solaire thermodynamique, aussi appelée énergie solaire concentrée (CSP : Concentrated Solar Power), consiste à concentrer les rayons du soleil à l'aide de miroirs ou de lentilles pour produire de la chaleur très intense. Cette chaleur est ensuite utilisée pour produire de l'électricité, souvent par le biais d'une turbine à vapeur.

#### 2.6.5.4.1. Fonctionnement du l'énergie solaire thermodynamique

Le solaire thermodynamique fait référence à l'ensemble des méthodes de conversion des rayons solaires en chaleur. En utilisant un système optique de concentration de la lumière solaire, on peut porter à des températures très élevées divers fluides. Après sa production, cette chaleur peut servir à générer de l'hydrogène ou à produire de l'énergie électrique. En effet, une centrale solaire thermodynamique a la capacité de convertir les rayons solaires en électricité. En résumé, elle absorbe le rayonnement solaire qu'un dispositif transforme en chaleur : cette chaleur est par la suite transformée en électricité. On peut emmagasiner la chaleur générée à l'aide d'un dispositif de stockage thermique. Donc, la génération d'électricité continue sans interruption même en cas de conditions météorologiques défavorables [24].

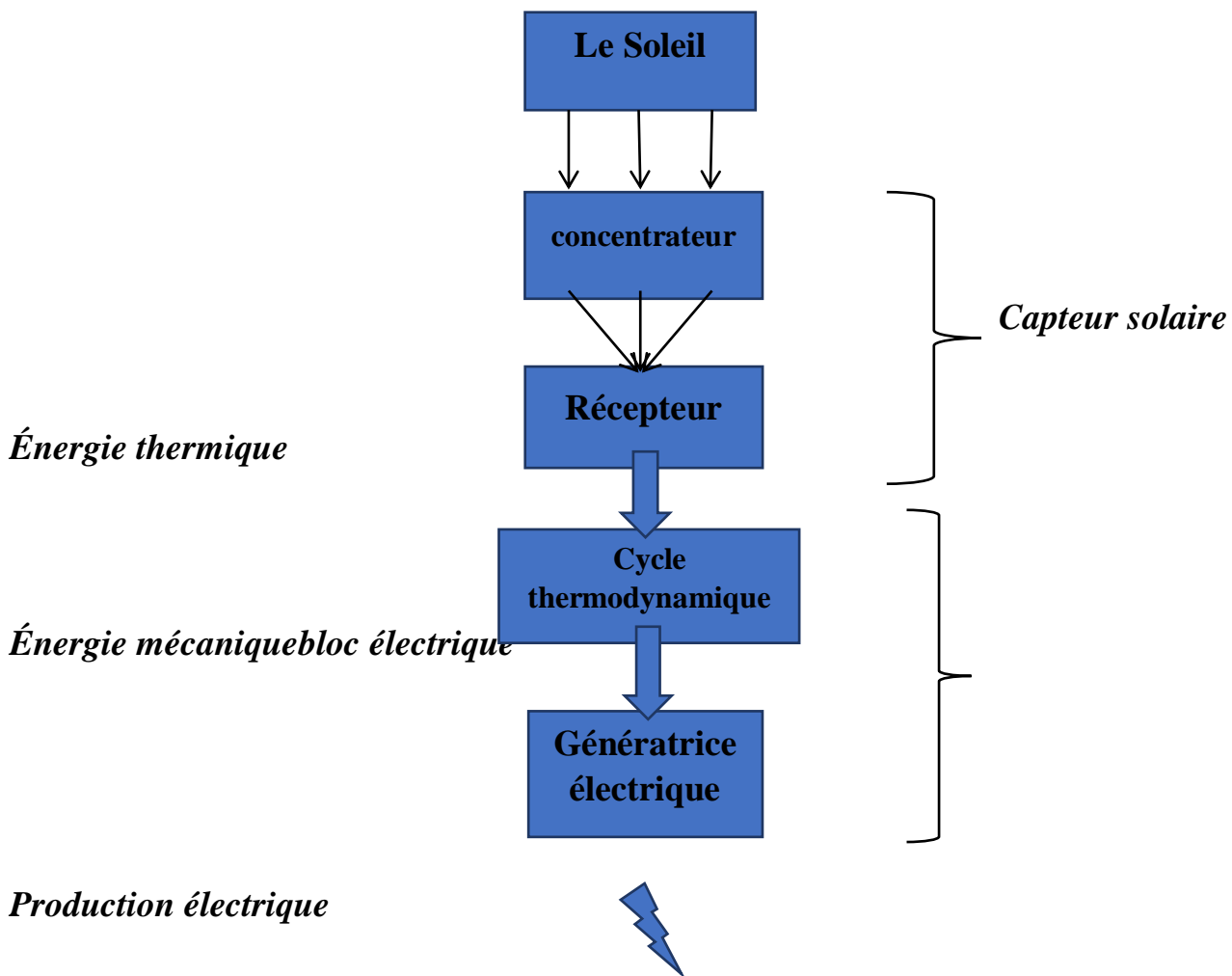


Figure 1.6 : Diagramme : fonctionnement de l'énergie solaire thermodynamique.

### 2.6.6. Les principaux avantages de l'énergie solaire

La popularité de l'énergie solaire peut être attribuée à ses nombreux bénéfices. C'est avant tout une source d'énergie renouvelable. En effet, elle se fonde sur l'énergie solaire, qui est à la fois riche, inlassable et durable. Elle est aussi disponible pour tous, offrant une indépendance relative face aux producteurs et fournisseurs d'énergie.

L'énergie solaire est aussi verte, contrairement aux énergies fossiles, car elle n'émet pas de gaz à effet de serre. Elle ne produit pas non plus de nuisance sonore, car les technologies qui permettent de convertir l'énergie solaire en électricité ou en chaleur sont silencieuses.

Finalement, une large partie des panneaux solaires peut être recyclée. Environ 94 % de leurs composants sont réutilisés d'après Energie, ce qui en fait une solution respectueuse de l'environnement sur le long terme [25].

### 2.6.7. L'avenir de l'énergie solaire

Le solaire pourrait devenir la première source d'énergie pour la production d'électricité. Contrairement à certaines idées reçues sur l'énergie solaire, cette dernière est déjà viable et compétitive. Elle se développe à grande vitesse partout dans le monde mais également en France même si l'hexagone a encore beaucoup de marge de progression en termes de production. Ce qui est certain, c'est que le soleil qui brille pour toutes les populations de la terre de la même manière, est une source inépuisable d'énergie. Chaque jour, le soleil fournit à la terre 3 kilowattheures par m<sup>2</sup> en moyenne. Une ressource que l'on est encore loin d'exploiter au maximum de ses capacités. Toutefois, le contexte géopolitique et économique mondial fait progresser à grands pas non seulement la recherche mais également les mentalités sur la question des énergies renouvelables et notamment le solaire. Le futur sera certainement davantage axé sur la production d'énergie solaire pour un certain nombre de raisons que nous aborderons dans les lignes qui suivent.

En effet, les politiques nationales ou Européennes accélèrent aujourd'hui les mesures de développement de production d'énergies renouvelables et le solaire en fait partie. D'autre part, la recherche progresse et fait bouger les lignes notamment en termes de coût et de capacité de production. Autant d'éléments qui nous amènent à conclure que le futur sera solaire [26].

## 3. Conclusion

Dans ce chapitre, nous sommes arrivés à la conclusion que Les énergies renouvelables représentent donc une option de substitution aux énergies fossiles pour plusieurs raisons : Elles sont habituellement moins nuisibles pour l'environnement, elles ne dégagent pas de gaz à effet de serre et ne génèrent pas de déchets. Elles sont illimitées et permettent une production décentralisée qui répond à la fois aux ressources et aux besoins locaux.

Finalement, elles procurent une autonomie énergétique considérable. L'augmentation du coût du pétrole, sa diminution, les conséquences environnementales des énergies fossiles et les récentes préoccupations concernant les centrales nucléaires anticipent une expansion des secteurs d'énergies renouvelables. Cette projection est basée sur la tendance actuelle, aussi bien dans les pays développés que dans les pays à croissance rapide.

# CHAPITRE 02

## L'Intelligence Artificielle appliquée à l'Optimisation des Systèmes d'Énergies Renouvelables

## 1. Introduction

Dans un contexte mondial marqué par la transition énergétique, l'intégration massive des énergies renouvelables dans les réseaux électriques pose de nouveaux défis en matière de gestion, de planification et d'optimisation. L'intelligence artificielle (IA) s'impose comme un levier puissant pour répondre à ces enjeux, notamment grâce à ses capacités d'analyse prédictive, d'apprentissage adaptatif et d'automatisation.

Ce chapitre a pour objectif de présenter les fondements de l'intelligence artificielle appliquée à l'optimisation des systèmes d'énergie renouvelable. Nous commencerons par une exploration des techniques de prédiction, en mettant en lumière leurs types, leurs avantages, ainsi que leur lien étroit avec les séries temporelles. Ensuite, nous aborderons les différentes formes d'apprentissage automatique, des méthodes supervisées aux approches auto-supervisées. Enfin, nous décrirons les grandes familles de l'IA utilisées dans ce domaine, en insistant sur les réseaux de neurones, le machine Learning et l'apprentissage profond, tout en introduisant des architectures avancées telles que LSTM et BILSTM.

Cette progression vise à doter le lecteur d'une compréhension claire et structurée des méthodes d'IA pertinentes pour anticiper, contrôler et optimiser la production, la consommation et le stockage d'énergie renouvelable.

## 2. Techniques de prédiction

### 2.1. Définition

Dans un contexte où les données constituent une ressource stratégique de premier ordre, la capacité à prédire des événements futurs à partir d'observations historiques s'impose comme un enjeu majeur dans de nombreux secteurs d'activité. Les techniques de prédiction, également appelées modèles prédictifs, jouent un rôle central dans des applications variées telles que la prévision des ventes, la détection précoce de maladies, l'estimation de la consommation énergétique, l'anticipation des défaillances industrielles ou encore l'analyse comportementale des utilisateurs.

De manière générale, la prédiction consiste à estimer une valeur inconnue – qu'elle soit continue (régression) ou catégorielle (classification) – à partir d'un ensemble de données d'entrée observées. Ce processus repose sur l'identification de relations sous-jacentes entre les variables, que l'on cherche à modéliser par des méthodes statistiques ou d'apprentissage automatique.

Quelques exemples concrets illustrent l'omniprésence de la prédiction :

- En **finance**, un analyste boursier tente de prédire l'évolution du cours d'un actif à partir de tendances passées.
- En **santé**, un modèle prédictif peut évaluer la probabilité qu'un patient développe une pathologie donnée en fonction de ses antécédents médicaux.
- En **marketing**, on cherche à anticiper le profil socio-comportemental d'un consommateur afin de personnaliser les offres qui lui sont proposées.

Sur le plan mathématique, un modèle de prédiction linéaire simple peut être exprimé par l'équation suivante :

$$\hat{y} = a + b \cdot x \quad (2.1)$$

Où :

$\hat{y}$  : représente la valeur prédite

X : désigne la variable d'entrée (ou caractéristique).

a et b sont des paramètres (ou coefficients) appris par le modèle lors de la phase d'entraînement.

## 2.2. Types de prédiction

La régression est une méthode d'apprentissage supervisé couramment employée afin de modéliser la relation entre une variable dépendante (ou cible) et une ou plusieurs variables indépendantes (ou caractéristiques). Divers types de régression sont disponibles, chacun étant approprié à des contextes particuliers. En fonction de la nature des données et de la forme de la relation entre les variables, divers types de régression peuvent être employés. Parmi les plus fréquents, on peut citer :

### 2.2.1. Prédiction numérique (régression)

La régression en apprentissage automatique est une méthode statistique employée afin de modéliser et d'analyser la relation entre une variable cible et une ou plusieurs variables explicatives. Son opération repose sur la calibration d'un algorithme mathématique afin de prédire une valeur numérique continue à partir de données observées. Ainsi, l'objectif principal consiste à fournir une estimation précise d'une variable continue en s'appuyant sur des données préexistantes. Diverses formes de prédiction numérique, telles que la régression, sont disponibles. :

- **La régression linéaire** : elle suppose une relation linéaire entre la variable cible et les variables explicatives. Ce modèle repose sur des principes simples tels que la minimisation de l'erreur quadratique moyenne.
- **La régression multiple** : contrairement à la régression linéaire simple, elle prend en compte plusieurs variables explicatives simultanément, permettant ainsi de modéliser des situations plus complexes.
- **La régression logistique** : bien qu'orientée vers des tâches de classification, elle utilise une fonction sigmoïde pour prédire la probabilité qu'un événement se produise, ce qui en fait une méthode adaptée à la classification binaire [27].

### 2.2.2. Prédiction catégorielle (classification)

La prédiction catégorielle, communément appelée classification, représente une tâche d'apprentissage supervisé dans les domaines de l'intelligence artificielle et de l'apprentissage automatique. Cette tâche implique la prédiction d'une étiquette ou catégorie en se basant sur des données d'entrée. Contrairement à la régression, où la sortie est une valeur continue, la classification produit une valeur discrète, appartenant à un ensemble fini de classes. Quelques exemples illustratifs incluent :

- **Lediagnostic médical**, où les classes peuvent être : malade / non malade ;
- **Larecognition d'image**, avec des classes telles que : voiture / piéton / feu rouge ;

- **La classification de texte**, notamment dans les filtres anti-spam : spam / non-spam.

Les types de classification se déclinent en plusieurs catégories :

- **Classification binaire** : où deux classes sont possibles (ex. : oui / non) ;
- **Classification multi classes** : impliquant plus de deux classes (ex. : rouge / vert / bleu) ;
- **Classification multi label** : où plusieurs classes peuvent être vraies simultanément, comme dans le cas d'une image contenant à la fois un chien et une voiture [28].

### 2.3. Avantages de la prédiction dans le domaine des énergies

La prédiction de la demande énergétique présente de nombreux avantages stratégiques pour les entreprises de services publics. Elle leur permet notamment de mieux se planifier en disposant d'une vision anticipée de la consommation ou de la charge future. Cette capacité d'anticipation contribue à la réduction des risques, en facilitant la prise de décisions économiquement viables concernant la production d'énergie à long terme ainsi que les investissements dans les infrastructures de transmission [29].

Elle offre également un soutien précieux dans l'identification des ressources nécessaires, telles que les carburants indispensables au fonctionnement des centrales, ainsi que les moyens logistiques nécessaires à la continuité de la production et à la distribution d'énergie à des prix compétitifs pour les consommateurs [29].

Les estimations de charge jouent un rôle essentiel dans la planification future en matière de volume, d'emplacement et de nature des unités de production. En identifiant les zones où la demande est élevée ou en croissance, les fournisseurs d'énergie peuvent localiser la production au plus proche des besoins, réduisant ainsi les pertes liées à la transmission et les coûts associés aux infrastructures de distribution [29].

Enfin, ces prévisions assistent les décideurs dans la planification de la maintenance des systèmes. Une bonne compréhension de la demande permet de déterminer les moments optimaux pour effectuer les opérations de maintenance, assurant ainsi un impact minimal sur les utilisateurs finaux [29].

### 2.4. Séries temporelles et leur rôle dans les systèmes énergétiques

Une série temporelle, aussi connue sous le nom de série chronologique, fait référence à une séquence de valeurs numériques illustrant la progression d'une variable particulière au cours du temps. Ces séquences de variables aléatoires peuvent être mathématiquement modélisées afin d'examiner leur comportement, tant pour analyser leur progression passée que pour anticiper leur évolution future. Cette modélisation est généralement basée sur des notions tirées de la probabilité et de la statistique [30].

On a parfois, à tort, tendance à considérer les séries temporelles comme un domaine exclusivement lié à l'économétrie. Cependant, cette discipline est assez récente comparée à l'utilisation des séries temporelles qui remontent à 1906 en astronomie et à 1968 en météorologie, domaines scientifiques dans lesquels elles étaient déjà largement employées [30].

Une série temporelle, également appelée série chronologique, désigne une suite de valeurs numériques représentant l'évolution d'une variable spécifique dans le temps. On peut modéliser mathématiquement ces séries de variables

aléatoires pour étudier leur comportement, aussi bien dans l'analyse de leurs tendances passées que dans la prévision de leur développement futur. Généralement, cette modélisation s'appuie sur des concepts issus de la statistique et de la probabilité [30].

Il arrive parfois, de manière erronée, de percevoir l'analyse des séries temporelles comme étant uniquement liée à l'économétrie. Toutefois, il est à noter que cette discipline est relativement nouvelle par rapport à l'usage des séries temporelles qui date de 1906 en astronomie et de 1968 en météorologie, deux champs scientifiques où elles étaient déjà largement utilisées [30].

### **3. Apprentissage automatique (Machine Learning)**

L'apprentissage constitue une étape cruciale dans le développement des réseaux de neurones, au cours de laquelle le comportement du réseau est modifié de manière itérative jusqu'à atteindre les performances souhaitées. Ce processus s'effectue par l'ajustement des poids (c'est-à-dire les connexions ou synapses) des neurones, en fonction d'informations sources clairement définies. L'apprentissage implique également l'extraction des relations cohérentes au sein des données utilisées pour entraîner le réseau.

L'objectif fondamental de l'apprentissage est de résoudre des problèmes variés tels que la prédiction ou la classification. Dans le cas des réseaux de neurones artificiels, l'apprentissage peut être considéré comme le processus de mise à jour des poids afin d'ajuster la réponse du réseau aux expériences et aux exemples présentés.

Les règles d'apprentissage sont nombreuses et peuvent être classées en deux grandes catégories : l'apprentissage supervisé et l'apprentissage non supervisé.

#### **3.1. Les types d'apprentissage**

##### **3.1.1. Apprentissage supervisé**

L'apprentissage supervisé est une technique d'apprentissage automatique où un algorithme se forme à partir d'un jeu de données labellisées, c'est-à-dire que chaque instance de formation correspond à une sortie spécifique ou à une catégorie. Le modèle est formé pour établir un lien entre les entrées et les sorties, afin de pouvoir réaliser des prédictions précises sur des données inédites. On utilise souvent cette méthode pour des tâches comme la classification, la régression ou la reconnaissance de formes [31].

Les exemples annotés servent de base à l'apprentissage :

- Classification d'images (ex. : reconnaissance de chiffres manuscrits avec MNIST) : l'ensemble de données MNIST contient des images de chiffres manuscrits, chacune associée à une étiquette (0 à 9). Un algorithme supervisé peut apprendre à classer correctement ces images [32].
- Prédiction de prix immobiliers (régression) : à partir de données telles que la superficie, le nombre de pièces ou encore la localisation, un modèle supervisé peut apprendre à prédire le prix d'un bien immobilier [33].
- Filtrage de spams dans les e-mails : des algorithmes supervisés, tels que les forêts aléatoires ou les machines à vecteurs de support (SVM), sont couramment utilisés pour classer les courriels en tant que « spam » ou « non-spam » à partir de caractéristiques comme le contenu, l'objet ou l'expéditeur [34].

### 3.1.2. Apprentissage non supervisé

Apprentissage non-supervisé est une stratégie d'apprentissage automatique où un algorithme se forme à partir de données non labellisées, c'est-à-dire sans indications précises de résultats, de classes ou de réponses. Le but est d'explorer des structures sous-jacentes, des agrégations ou des schémas dissimulés dans les données. On utilise particulièrement cette méthode dans des secteurs tels que l'analyse de données, la réduction de la dimensionnalité ou encore la détection d'anomalies [35].

Quelques exemples d'application incluent :

- Segmentation de clientèle : les entreprises recourent à des techniques comme le clustering (ex. : k-means) pour regrouper les clients selon des comportements d'achat similaires, facilitant ainsi les stratégies de marketing ciblé [36].
- Compression de données : l'usage d'auto-encodeurs permet de réduire la dimensionnalité tout en conservant l'information essentielle, ce qui est particulièrement utile pour le stockage ou le traitement efficace de grandes bases de données [37].
- Détection d'anomalies : l'apprentissage non supervisé peut être utilisé pour identifier des valeurs aberrantes (ex. : fraudes bancaires) en détectant les données qui s'écartent significativement des structures apprises, comme c'est le cas avec les modèles de forêts d'isolation [38].

### 3.1.3. Apprentissage semi-supervisé

L'apprentissage semi-supervisé est une méthode mixte d'apprentissage automatique qui fusionne un petit échantillon de données labellisées avec une vaste quantité de données non labellisées pour optimiser l'efficacité des modèles prédictifs. Ce modèle se fonde sur l'idée que la structure inhérente des données peut être utilisée pour transmettre les informations des étiquettes existantes vers les exemples sans étiquette. Cette méthode se révèle particulièrement appropriée dans les situations où l'annotation manuelle des données est onéreuse ou complexe à mettre en œuvre [39].

L'apprentissage semi-supervisé est appliqué dans divers domaines, parmi lesquels :

- Classification d'e-mails comme spam ou non-spam : lorsqu'une entreprise dispose d'un petit jeu de courriels classifiés et d'un grand ensemble non étiqueté, un algorithme semi-supervisé permet d'améliorer significativement la précision du filtrage des courriels indésirables [40].
- Reconnaissance faciale : dans les systèmes de reconnaissance, seuls quelques visages peuvent être associés à une identité précise. Un algorithme semi-supervisé peut exploiter un grand nombre d'images non étiquetées pour affiner le modèle de reconnaissance [41].
- Analyse de sentiments sur les réseaux sociaux : à partir d'un petit nombre de messages annotés (ex. : positif, négatif, neutre), les modèles semi-supervisés permettent d'étendre l'analyse à un volume important de données non étiquetées, améliorant ainsi la portée des analyses d'opinion [42].

### 3.1.4. Apprentissage auto-supervisé

L'apprentissage auto-supervisé représente une méthode d'apprentissage automatique où les algorithmes font appel à des données non labellisées pour créer leurs propres étiquettes en se basant sur la structure inhérente de ces données. Cette approche s'appuie sur l'accomplissement de tâches prétextes (comme la prédiction d'un segment absent d'une image ou d'une séquence), facilitant le pré-entraînement de modèles sur une vaste quantité de données non étiquetées. On peut par la suite perfectionner ces modèles en utilisant des données annotées. Cette méthode a récemment captivé l'attention, surtout dans les secteurs de la manipulation d'images et du traitement de langage naturel, car elle offre une solution aux limitations associées à l'accessibilité restreinte des données annotées [43].

Quelques exemples d'application de l'apprentissage auto-supervisé incluent :

- Pré-entraînement de modèles de traitement du langage naturel (ex. : BERT) : le modèle BERT (*Bidirectional Encoder Representations from Transformers*) est préentraîné via une tâche de masquage de mots, consistant à prédire des mots masqués dans une phrase à partir du contexte environnant [44].
- Représentation d'images avec SimCLR : le modèle SimCLR apprend des représentations visuelles en maximisant la similarité entre différentes augmentations d'une même image, sans avoir recours à des étiquettes explicites. Cette tâche contrastive constitue une forme d'auto-supervision [45].
- Apprentissage de représentations audio avec wav2vec : le modèle wav2vec 2.0 est préentraîné de manière auto-supervisée en apprenant à prédire des représentations de segments audio masqués à partir du contexte, permettant ainsi la création de systèmes de reconnaissance vocale efficaces sans données annotées [46].

## 4. Intelligence artificielle appliquée à l'énergie renouvelable

L'IA trouve ses origines en 1943, avec la parution de l'article *A Logical Calculus of Ideas Immanent in Nervous Activity* par Warren McCulloch et Walter Pitts. Cette étude présente le premier modèle mathématique conçu pour la création d'un réseau de neurones. Snarc, le premier ordinateur basé sur un réseau de neurones, a été conçu en 1950 par Marvin Minsky et Dean Edmonds, deux étudiants de l'université Harvard. Cette même année, Alan Turing suggère le Test de Turing, qui reste encore à ce jour un instrument de référence pour l'évaluation des systèmes d'IA. Ce test établit les principes fondamentaux de l'intelligence artificielle, en précisant ses buts essentiels : reproduire ou simuler l'intelligence humaine à travers des machines.

Grâce à l'intelligence artificielle, les ordinateurs peuvent apprendre de données complexes et résoudre des problèmes pratiques dans divers secteurs, y compris la santé, atteignant parfois des niveaux de performance comparables ou même supérieurs à ceux d'experts humains. Elle fait référence à toutes les techniques informatiques d'analyse de données où les algorithmes apprennent principalement par la pratique, sans être spécifiquement codés par des experts. Dans cet article, nous employons le terme IA de manière exhaustive, englobant aussi bien les méthodes traditionnelles d'apprentissage automatique (AM) que celles d'apprentissage profond (AP) [47].

Les méthodes d'AM conventionnelles, qui ne requièrent pas d'équipement spécifique, sont employées depuis de nombreuses années dans le secteur médical, particulièrement dans des domaines tels que l'hépatologie et la gastro-entérologie [48].

Dans le domaine de l'analyse d'images radiologiques, les méthodes traditionnelles basées sur des caractéristiques élaborées manuellement sont appelées radio Mique. En complément de ces techniques bien établies, l'apprentissage profond (DL) a progressé au cours de la dernière décennie, stimulé par les avancées algorithmiques, le perfectionnement des équipements informatiques et l'accessibilité croissante de vastes ensembles de données. Les méthodes de DL, bien que conceptuellement similaires aux techniques traditionnelles de ML, se caractérisent par une quantité significativement plus élevée de paramètres libres. Cela les rend plus flexibles et mieux appropriées pour traiter des données complexes comme les images médicales ou les séquences linguistiques.

Dans le domaine médical, les techniques DL les plus courantes comprennent les réseaux de neurones artificiels (employés notamment pour le traitement d'images [49] et l'analyse de séries temporelles) ainsi que les transformateurs, qui sont désormais utilisés autant pour l'analyse d'images que pour le traitement du langage. À l'opposé des approches classiques, les modèles d'apprentissage profond n'ont pas besoin de déterminer manuellement les caractéristiques : ils sont capables d'apprendre de manière autonome à extraire les représentations appropriées à partir des données, selon un critère d'appréciation, comme le résultat clinique. Les techniques de Deep Learning surpassent souvent les méthodes traditionnelles grâce aux outils technologiques modernes, et jouent un rôle prépondérant dans l'application de l'IA à l'hépatologie. Cependant, la séparation entre ces deux méthodes n'est pas stricte : de nombreuses recherches combinent des caractéristiques dérivées du DL avec des caractéristiques élaborées manuellement [50].

Sur le plan opérationnel, les techniques ML/DL peuvent jouer deux rôles essentiels : d'un côté, elles permettent d'automatiser l'analyse des données, une fonction traditionnellement dévolue aux spécialistes humains ; de l'autre côté, elles sont capables de déceler des traits délicats dans des ensembles de données complexes, souvent hors de portée pour une analyse visuelle directe [51].

#### **4.1. Définition générale de l'IA**

L'intelligence artificielle (IA) englobe diverses méthodes et techniques qui autorisent les machines à reproduire certaines capacités de l'intelligence humaine, comme l'acquisition de connaissances, le processus de réflexion, la prise de décisions ou encore la compréhension du langage naturel. Cette discipline se compose de diverses catégories majeures, chacune étant fondée sur des principes distincts. L'une de ces techniques, l'apprentissage automatique (Machine Learning), est particulièrement centrale : elle permet aux systèmes informatiques de déduire des schémas à partir des données, sans qu'il soit indispensable de spécifier explicitement les directives de fonctionnement. L'Apprentissage Profond (Deep Learning), pour sa part, représente une ramification sophistiquée de l'apprentissage automatique, basée sur l'emploi de réseaux neuronaux artificiels profonds, particulièrement efficaces pour gérer d'importants ensembles de données complexes et non organisées. La figure 2.1 Présentela hiérarchie des domaines de l'Intelligence Artificielle.

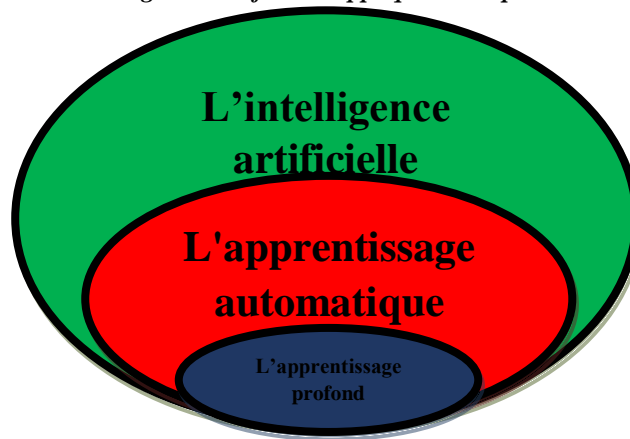


Figure 2.1 : Hiérarchie des domaines de l'Intelligence Artificielle

### 4.2.1. Réseaux de neurones artificiels

On peut définir un réseau de neurones comme un graphe orienté, constitué d'un ensemble d'unités de traitement de base, disposées en couches successives reliées entre elles. Ces éléments, couramment désignés comme neurones artificiels, sont connectés par des associations organisées, facilitant le transfert d'informations conformément à une architecture préétablie. Un réseau neuronal artificiel, malgré son inspiration tirée de la neurobiologie, est principalement un modèle mathématique de ce graphe qui vise à simuler certaines fonctions cognitives détectées dans les systèmes biologiques.

Ces réseaux sont élaborés pour reproduire certaines caractéristiques essentielles du traitement neuronal naturel, comme la faculté d'apprendre, le stockage d'informations et la résilience face à des données partielles ou perturbées. Depuis l'adoption de cette méthode, une multitude de modèles neuronaux ont vu le jour, tous basés sur des concepts issus du connexionnisme, une perspective qui postule que l'intelligence découle d'un vaste réseau d'unités élémentaires interconnectées.

Ainsi, les modèles de connexionnisme artificiel se fondent sur des analogies biologiques concernant la structure et le comportement général. Peu importe le modèle choisi, les principes de base d'interconnexion entre les unités et de traitement simultané de l'information sont toujours gardés. En se fondant sur ces principes, les réseaux de neurones se sont révélés être des instruments efficaces dans divers domaines d'application, démontrant leur efficacité en matière d'apprentissage automatique et de traitement de données complexes.

#### 4.2.1.1. Perceptron multicouche (MLP)

Dans le domaine de l'intelligence artificielle, plus précisément dans le champ de l'apprentissage automatique, le perceptron multicouche (Multilayer Perceptron – MLP) est une forme de réseau de neurones artificiel organisé en plusieurs strates hiérarchisées. Ce modèle est constitué d'au moins trois strates distinctes : une couche d'entrée, une ou plusieurs couches intermédiaires, et enfin, une couche de sortie. Chaque couche contient un nombre spécifique de neurones, qui peut différer d'une couche à l'autre. Dans cette structure, le transfert d'informations se fait de manière unidirectionnelle, c'est-à-dire de la couche d'entrée à la couche de sortie, sans rétroaction : ceci est connu sous le nom de réseau neuronal à propagation directe (feedforward neural network). Les neurones situés dans la couche de sortie délivrent la réponse définitive du système aux données d'entrée [52].

Frank Rosenblatt a introduit le perceptron, dans sa forme originale, en 1957 au Laboratoire Aéronautique de Cornell. Initialement, ce modèle possédait seulement une couche de traitement unique et une sortie exclusive, à laquelle toutes les entrées étaient directement rattachées [53]. Voici une présentation de l'architecture du perceptron multicouche, soulignant les divers éléments structuraux et opérationnels de ce modèle neuronal.

La **Figure 2.2** illustre l'architecture typique d'un perceptron multicouche. On y distingue clairement les trois types de couches :

- La couche d'entrée, qui reçoit les données initiales,
- Une ou plusieurs couches cachées, où s'opèrent les transformations internes non observables,
- Et la couche de sortie, qui génère les résultats finaux du réseau.

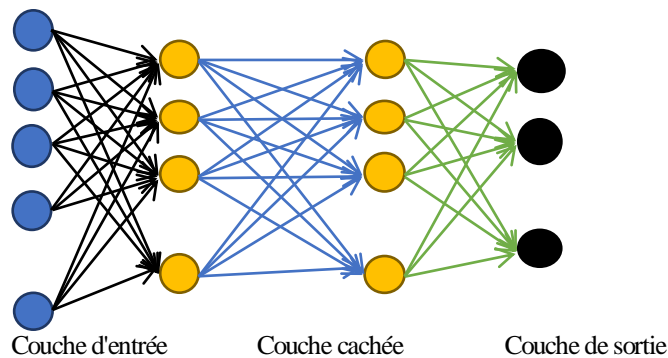


Figure 2.2 : Architecture typique d'un perceptron multicouche

- La couche d'entrée : Il s'agit de la première couche du réseau. Elle reçoit directement les données d'entrée, c'est-à-dire les variables explicatives du problème à modéliser. La taille de cette couche est déterminée par le nombre de caractéristiques (ou variables) présentes dans les données.
- La ou les couches cachées : Situées entre la couche d'entrée et la couche de sortie, ces couches constituent le cœur du traitement effectué par le réseau. Un réseau peut comporter une ou plusieurs couches cachées. Lorsque la fonction de transfert est linéaire, les couches cachées peuvent être assimilées à une transformation linéaire unique. Toutefois, dans la pratique, on utilise généralement des fonctions d'activation non linéaires (telles que ReLU, sigmoid, tanh), ce qui confère au réseau sa capacité à modéliser des relations complexes. Le nombre de neurones dans les couches cachées n'est pas fixé a priori ; il constitue un hyper paramètre qui doit être ajusté empiriquement, en fonction des performances du modèle sur un ensemble de validation.
- La couche de sortie : Elle représente la dernière couche du réseau, produisant la réponse finale du modèle à partir des traitements effectués dans les couches précédentes. Sa taille dépend directement du nombre de variables cibles à prédire (par exemple : une seule sortie pour une tâche de régression uni-variée, ou plusieurs pour une classification multi-classe).

En complément de cette architecture, le MLP (Multilayer Perceptron) implique deux types de paramètres :

#### a) Paramètres appris :

- Les poids synaptiques, qui définissent les connexions entre neurones, sont mis à jour de manière itérative au cours de l'apprentissage, afin de minimiser l'erreur du modèle.

- Les biais, associés à chaque neurone, permettent de décaler la fonction d'activation, jouant un rôle important dans l'ajustement de la sortie neuronale.

#### b) Hyper paramètres (définis manuellement par l'utilisateur) :

Il s'agit de paramètres fixés avant l'entraînement du modèle, tels que le taux d'apprentissage, le nombre d'itérations (ou d'époques), le nombre de neurones dans les couches cachées, ou encore la fonction d'activation choisie. Ces hyper paramètres influencent fortement les performances et la capacité de généralisation du réseau [54], [55].

Le tableau suivant présente les principaux hyper paramètres du MLP, accompagnés de leur description :

Tableau 2.1 Hyper paramètres de MLP

Hyper paramètre	Description
Nombre de couches cachées	Plus il y a de couches, plus le modèle peut apprendre des représentations complexes
Nombre de neurones par couche	Plus de neurones = plus de puissance de calcul
Fonction d'activation	Elle décide si un neurone "s'active"
Taux d'apprentissage (Learning rate)	Contrôle la vitesse d'ajustement des poids pendant l'entraînement
Fonction de perte (lossfunction)	Dépend du type de tâche.
Nombre d'époques (epochs)	Combien de fois on passe sur tout le dataset pendant l'entraînement
Taille du batch (batch size)	Combien d'exemples sont traités en une seule itération.
Optimiseur (optimizer)	Méthode pour ajuster les poids

### 4.2.1.2. Réseau à fonctions de base radiale (RBF)

#### 4.2.1.2.1. Définition

L'architecture de réseaux neuronaux en couches appelée réseaux à fonctions de base radiale (RBF) est dérivée de la technique d'interpolation utilisant des fonctions de base radiale. Broomhead et Lowe ont introduit cette méthode pour la première fois dans le contexte des réseaux inspirés du fonctionnement du cerveau en 1988, avant que Moody et Darken ne présentent, en 1989, le premier algorithme d'apprentissage lié à ce genre de réseaux. Depuis, le réseau RBF est perçu comme une option encourageante par rapport au perceptron multicouche, offrant généralement une convergence plus rapide et une meilleure performance dans de nombreux scénarios.

Tout d'abord destiné à des missions de classification, le réseau RBF a prouvé sa flexibilité et a été utilisé avec succès dans différents secteurs, comme l'approximation de fonctions, la reconnaissance vocale, ainsi que d'autres défis relatifs à la modélisation et à l'analyse de données [56].

Un Réseau de Fonctions Radiales de Base (RBF) est généralement constitué de trois couches distinctes : une couche d'entrée, une couche cachée constituée de neurones qui mettent en œuvre des fonctions noyaux (souvent des fonctions gaussiennes), et une couche de sortie, généralement équipée de neurones employant une fonction d'activation linéaire.

Les neurones de la couche cachée évaluent la distance entre les données d'entrée du réseau et leur point central (ou prototype), puis utilisent une fonction d'activation de type gaussien pour générer une réponse proportionnelle à cette distance

La sortie globale du réseau résulte ensuite d'une combinaison linéaire pondérée des activations issues des neurones de la couche cachée, où chaque contribution est modulée par un poids synaptique spécifique.

La figure suivante illustre l'architecture générale d'un réseau RBF, mettant en évidence le rôle de chaque couche ainsi que le principe de fonctionnement des fonctions noyaux.

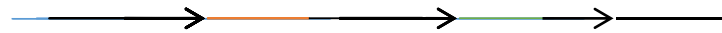


Figure 2.3 : Architecture générale d'un réseau RBF.

#### **4.2.2. Apprentissage Automatique (Machine Learning)**

Le Machine Learning (ML) représente un domaine crucial de l'intelligence artificielle, axé sur la création et l'examen d'algorithmes statistiques capables d'apprendre à partir de données et de se généraliser à des données inédites, ce qui permet d'accomplir des tâches sophistiquées sans directives spécifiques [57].

L'apprentissage profond (Deep Learning), un sous-domaine du ML, a bénéficié de progrès significatifs qui ont conduit les réseaux de neurones — une classe particulière d'algorithmes statistiques — à surpasser plusieurs méthodes traditionnelles en matière de performance et de capacité de généralisation [58].

Aujourd'hui, le machine learning est couramment employé dans divers domaines, y compris le traitement du langage naturel, la vision informatique, la reconnaissance de la parole, le filtrage des emails, et s'applique également dans des domaines comme l'agriculture et la médecine [59]. Quand on l'applique à des enjeux commerciaux ou industriels, cela devient une analyse prédictive, dont le but est de prévoir des événements futurs en se basant sur des données passées.

En théorie, l'apprentissage automatique s'appuie sur des techniques provenant de la statistique et de l'optimisation mathématique, en particulier la programmation mathématique. Cela est fortement associé à la fouille de données, un champ connexe dédié à l'analyse exploratoire des données (AED), généralement réalisé par le biais de méthodes d'apprentissage non supervisé [60].

#### 4.2.2.1. Machine à vecteurs de support (SVM)

Les machines à vecteurs de support, aussi appelées Support Vector Machines (SVM), sont des modèles supervisés d'apprentissage automatique particulièrement efficaces pour traiter les problèmes de classification et de régression. Leur but premier est de déterminer l'hyperplan idéal qui distinctement divise les diverses classes de données, en maximisant la distance entre cet hyperplan et les points de données adjacents, connus sous le nom de vecteurs de support. Les SVM, qui ont été introduites dans les années 1990, s'appuient sur la théorie d'apprentissage statistique développée par les informaticiens russes Vladimir Vapnik et Alexey Chervonenkis. Cette théorie est également connue sous le nom de théorie de Vapnik-Chervonenkis. L'association de ce cadre théorique strict et de leur aptitude à manipuler des données de haute dimension a conduit à une adoption rapide de ces modèles, qui ont fait preuve d'une efficacité remarquable. en pratique [61].

Le Principe de fonctionnement d'une Machine à Vecteurs de Support (SVM) est illustré sur la figure2.4

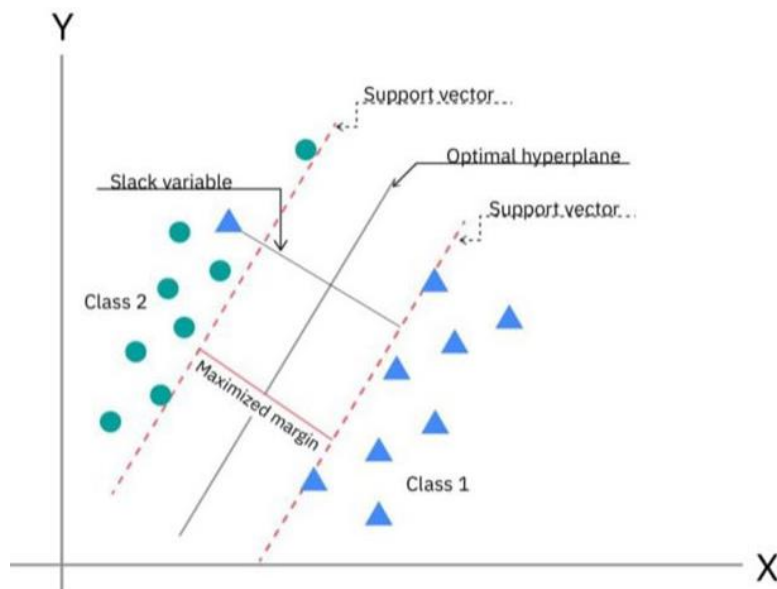


Figure2.4 : Principe de fonctionnement d'une Machine à Vecteurs de Support (SVM).

Les composants essentiels d'une SVM sont les suivants :

- Données d'entrée** : Chaque observation est représentée sous la forme d'un vecteur de caractéristiques dans un espace à  $n$  dimensions [62].
- Hyperplan de séparation** : Il s'agit de la frontière décisionnelle qui divise l'espace des caractéristiques afin de séparer les différentes classes. Cette frontière peut prendre la forme d'une ligne en deux dimensions, d'un plan en trois dimensions, etc. [63].

c) **Marge maximale** : La SVM cherche à maximiser la distance entre l'hyperplan et les vecteurs de support, c'est-à-dire les points de données les plus proches de l'hyperplan appartenant à chaque classe [64].

d) **Vecteurs de support** : Ces points de données déterminent la position et l'orientation de l'hyperplan optimal, jouant un rôle central dans la définition du modèle.

e) **Fonction noyau (Kernel)** : Lorsque les données ne sont pas séparables linéairement dans l'espace d'origine, la SVM applique une transformation via une fonction noyau qui projette les données dans un espace de dimension supérieure, facilitant ainsi la séparation linéaire des classes.

#### 4.2.2.2. Régression à vecteurs de support (SVR)

La régression par vecteurs de support (SVR) est une méthode d'apprentissage automatique spécifiquement conçue pour prédire des valeurs continues, en opposition aux machines à vecteurs de support (SVM) qui sont généralement employées pour des missions de classification. La SVR, qui est une variante des SVM, se fonde sur des principes analogues, en particulier l'application du concept de marge pour établir la meilleure approximation envisageable. Cette marge permet une certaine tolérance aux erreurs de prédiction tout en minimisant l'effet des valeurs extrêmes.

La SVR, bien qu'elle soit moins courante que les SVM, s'est avérée être un instrument efficace pour l'estimation de fonctions à valeurs réelles. En tant qu'approche d'apprentissage supervisé, elle se base sur une fonction de perte symétrique qui sanctionne de manière similaire les erreurs de prédiction dépassant la marge, qu'elles soient supérieures ou inférieures. Formulée sous la forme d'un problème d'optimisation, la SVR cherche à élaborer une fonction de perte convexe qui soit insensible aux petites fluctuations, puis à définir le « tube » de largeur minimale qui englobe la plupart des données d'apprentissage [65]. La figure 2.5 illustre le Principe de la Régression à Vecteurs de Support (SVR).

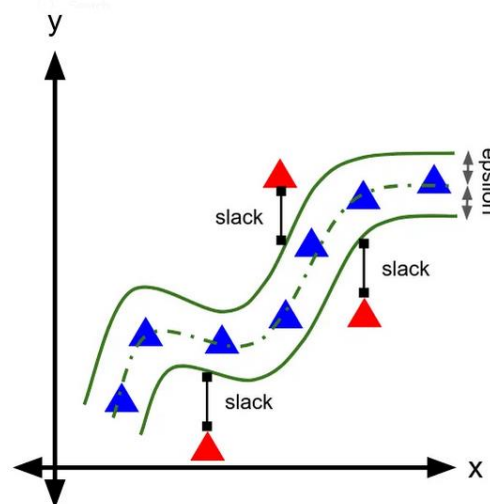


Figure 2.5 : Principe de la Régression à Vecteurs de Support (SVR).

#### 4.2.3. Apprentissage Profond (Deep Learning)

Le Deep Learning, ou apprentissage profond en français, est une spécialité de l'intelligence artificielle basée sur l'emploi de réseaux neuronaux artificiels à plusieurs strates hiérarchiques. Ces structures profondes facilitent le modelage et l'apprentissage de représentations complexes des données à divers niveaux d'abstraction, ce qui rend la

résolution de tâches avancées possible. Le Deep Learning a progressé de manière significative et rapide dans différents domaines, notamment dans l'analyse des signaux visuels et sonores. Il a particulièrement transformé des applications comme la reconnaissance de la parole, la détection faciale, l'analyse d'images et le traitement automatique du langage naturel. Ces avancées ont été grandement facilitées grâce à des investissements colossaux, publics et privés, en particulier de la part de grandes entreprises technologiques comme Google, Apple et Facebook. Amazon et Microsoft (GAFAM), au cours des années 2000 [66].

#### 4.2.3.1. Mémoire à long terme (LSTM)

Pour pallier les contraintes des réseaux neuronaux récurrents traditionnels, notamment les problèmes de disparition ou d'explosion du gradient lors de l'apprentissage de séquences étendues, Hochreiter et Schmidhuber (1997) ont proposé l'usage des réseaux à mémoire à long et court terme (LSTM). Les architectures LSTM sont une forme sophistiquée de RNN spécialement élaborée pour saisir les dépendances à long terme dans les données en séquence. Ils conviennent particulièrement bien aux missions de classification, de traitement et de prédiction sur des séries temporelles, où les événements significatifs peuvent être distants les uns des autres.

L'architecture LSTM s'articule principalement autour de l'introduction de blocs de mémoire, qui regroupent des cellules mémoires et un système de portes régulant le passage d'informations. Ces portes — d'entrée, de sortie et d'oubli — donnent au réseau la capacité de déterminer quelles informations doivent être retenues, éliminées ou transmises à l'étape suivante. Par conséquent, à l'inverse des RNN traditionnels, les LSTM ont la capacité de conserver une mémoire stable sur de longues séquences temporelles tout en autorisant une actualisation dynamique des états internes [68]. L'architecture typique d'un LSTM est présentée dans la figure 2.6

Chaque cellule mémoire comporte deux états :

- $h_t$  : l'état caché, représentant la mémoire à court terme,
- $c_t$  : l'état de la cellule, représentant la mémoire à long terme.



Figure 2.6 : Architecture typique d'un LSTM.

Les trois principales portes assurent un contrôle précis sur ces états :

**a) La porte d'entrée (*input gate*)**

Elle décide des nouvelles informations à stocker dans la mémoire à long terme :

$$f_t = \sigma(W_t \cdot [h_{t-1}, x_t] + b_t) \quad (2.2)$$

$$C_t = f_t \times C_{t-1} \quad (2.3)$$

Avec :

- $h_{t-1}$  : sortie précédente,
- $x_t$  : entrée actuelle,
- $b_t$  : biais,
- $W_t$  : poids,
- $\sigma$  : fonction sigmoïde.

**b) La porte d'oubli (*forgetgate*)**

Elle supprime les informations devenues obsolètes dans la cellule mémoire :

$$f_c = \tanh(W_c \cdot h_{t-1} + b_c) \quad (2.4)$$

$$I_t = \sigma(W_t \cdot [h_{t-1}, x_t] + b_i) \quad (2.5)$$

L'état actuel est mis à jour comme suit :

$$C_t = f_t \times C_{t-1} + f_c \times I_t \quad (2.6)$$

Avec :

- $\tanh$ : fonction d'activation tangente hyperbolique,
- $C$  : valeur candidate.

**c) La porte de sortie (*output gate*)**

Elle génère la sortie actuelle du LSTM :

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (2.7)$$

$$h_t = o_t \times \tanh(C_t) \quad (2.8)$$

Ces mécanismes combinés permettent aux réseaux LSTM de mémoriser, de sélectionner et d'omettre efficacement des informations à travers des séquences temporelles longues, tout en assurant la stabilité du processus d'apprentissage grâce à une rétropropagation plus robuste. Ce contrôle précis du flux d'information offre aux LSTM des performances supérieures dans diverses applications comme la traduction automatique, la reconnaissance vocale, l'analyse de séries chronologiques et la modélisation de dépendances temporelles complexes [69].

#### 4.2.3.2. Mémoire à long terme bidirectionnelle (BILSTM)

Le modèle Bidirectional Long Short-Term Memory (BILSTM) est une version étendue des réseaux de neurones à mémoire longue et courte (LSTM), qui permet l'utilisation conjointe d'informations issues du passé et du futur dans une séquence spécifique [70]. À la différence du LSTM traditionnel, qui analyse les données uniquement dans une direction temporelle (généralement du passé vers le futur), le BILSTM s'appuie sur deux couches LSTM en parallèle :

l'une se dirigeant vers l'avant (LSTM à sens unique) et l'autre vers l'arrière (LSTM inversé). Les résultats des deux trajectoires sont par la suite fusionnés — soit par concaténation, soit par addition — pour générer une représentation contextuelle améliorée de chaque composant de la séquence [71].

Cette double approche permet au modèle de déceler des caractéristiques complexes liées au temps en saisissant non seulement les dépendances du passé, mais aussi celles à venir. Ce contexte renforcé rend le BILSTM particulièrement performant dans des domaines tels que le traitement du langage naturel (NLP), la traduction automatique, l'analyse de sentiments, la détection d'entités nommées ou encore l'étude des séries temporelles.

Sur le plan computationnel, un réseau BILSTM calcule d'abord la séquence de sortie de la couche cachée avant  $h_{t \rightarrow}$ , puis celle de la couche arrière  $h_{t \leftarrow}$ , respectivement selon les équations suivantes :

$$h_t^{\rightarrow} = H(W_{xh}^{\rightarrow} X_t + W_{hh}^{\rightarrow} h_{t+1}^{\rightarrow} + b_h^{\rightarrow}) \quad (2.9)$$

$$h_t^{\leftarrow} = H(W_{xh}^{\leftarrow} X_t + W_{hh}^{\leftarrow} h_{t+1}^{\leftarrow} + b_h^{\leftarrow}) \quad (2.10)$$

$$y_t = W_{xy}^{\rightarrow} h_t^{\rightarrow} + W_{xy}^{\leftarrow} h_t^{\leftarrow} + b_y \quad (2.11)$$

Où :

H représente la fonction d'activation, souvent implémentée via la tangente hyperbolique ou la fonction ReLU, et  $X_t$  l'entrée à l'instant t. La sortie finale est obtenue à partir de la combinaison des deux états cachés  $h_{t \rightarrow}$  et  $h_{t \leftarrow}$ , selon :

$$y_t = \sigma(h_t^{\rightarrow}, h_t^{\leftarrow}) \quad (2.12)$$

L'architecture complète d'un modèle BILSTM, tel qu'illustré dans le schéma de la figure 2.7, comprend généralement six couches principales [72] :

1. **Couche d'entrée de séquence** : taille d'entrée fixée à 12.
2. **Couche BILSTM** : composée de deux sous-ensembles de cellules LSTM (forward et backward), chacun maintenant ses propres états internes. Nombre d'unités cachées : 150.
3. **Couche de dropout** : insérée entre les couches LSTM afin de prévenir le surapprentissage. Elle désactive aléatoirement certains neurones à chaque itération, avec une probabilité typique de 0,75.
4. **Couche entièrement connectée (dense)** : convertit les représentations latentes en prédictions. Nombre de classes de sortie : 15.
5. **Couche Softmax** : transforme les activations en probabilités de classes.
6. **Couche de classification** : effectue la prédiction finale en fonction de la classe ayant la probabilité la plus élevée [73]. La figure 2.7 présente l'architecture typique d'un BiLSTM.



Figure 2.7 : Architecture typique d'un BiLSTM

En résumé, l'approche bidirectionnelle du BiLSTM offre une compréhension contextuelle complète des séquences en exploitant les corrélations temporelles dans les deux directions. Cette capacité est essentielle dans de nombreux systèmes modernes d'intelligence artificielle traitant des données séquentielles complexes.

### 4.2.3.3. Comparaison entre LSTM et BILSTM

Afin de mieux cerner les spécificités et les cas d'usage de chaque architecture, le tableau ci-dessous présente une comparaison synthétique entre le LSTM classique et sa version bidirectionnelle, le BILSTM.

Tableau 2.2 : Comparaison entre LSTM et BILSTM [74].

Fonctionnalité	Mémoire à long terme (LSTM)	Mémoire à long terme bidirectionnelle (BILSTM)
<b>Directionnalité</b>	Traite les séquences dans une direction (vers l'avant ou vers l'arrière).	Traite les séquences dans les deux sens (avant et arrière).
<b>Informations contextuelles</b>	A accès uniquement au contexte passé lors de la prédiction de l'état suivant.	Accède au contexte passé et futur, offrant une compréhension plus complète de la séquence.
<b>Complexité</b>	Plus simple en architecture et en calcul.	Plus complexe en raison du traitement en double séquence, nécessitant davantage de ressources de calcul.
<b>Performance</b>	Suffisant pour les tâches où le contexte passé est adéquat.	Fournit généralement de meilleures performances dans les tâches où le contexte de séquence complet (passé et futur) est important.
<b>Cas d'utilisation</b>	Idéal pour la prédiction de séries chronologiques, la reconnaissance vocale et d'autres tâches nécessitant uniquement un contexte passé.	Idéal pour les tâches PNL telles que l'analyse des sentiments, la traduction automatique et la reconnaissance d'entités nommées, où la compréhension de l'ensemble du contexte de la séquence est bénéfique.
<b>Exigences de mémoire</b>	Nécessite généralement moins de mémoire.	Nécessite plus de mémoire en raison de paramètres et de traitements supplémentaires.
<b>Temps de formation</b>	Plus rapide grâce au traitement unidirectionnel.	Plus lent en raison du traitement des séquences dans les deux sens.

## 5. Conclusion.

Pour conclure, ce chapitre a permis de poser les fondements théoriques nécessaires à la compréhension de la prédiction et de son rôle stratégique, notamment dans le secteur énergétique. En exposant les différentes catégories de prédiction, nous avons souligné son importance dans l'anticipation des comportements futurs liés à la production et à la consommation d'énergie, contribuant ainsi à une gestion plus efficace et durable des ressources.

Nous avons ensuite approfondi la notion d'intelligence artificielle, en présentant ses principales déclinaisons et en mettant en évidence les avancées majeures qui ont favorisé son essor. L'accent a été mis sur les méthodes de prédiction issues de l'IA, capables de tirer parti des données historiques pour générer des estimations robustes et pertinentes. Ces techniques s'avèrent particulièrement adaptées à notre projet, qui repose sur une modélisation intelligente des phénomènes énergétiques.

# **CHAPITRE03**

Méthodologie et Outils pour le  
développement de l'Interface  
MATLAB - Etude de Cas.

## 1. Introduction

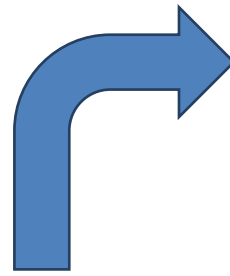
Ce chapitre présente les résultats obtenus dans le cadre de notre projet visant à concevoir une interface utilisateur sous MATLAB dédiée à la planification et à l'installation de sites d'énergies renouvelables. L'objectif principal était de développer un outil interactif permettant de simuler et d'optimiser l'implantation de systèmes énergétiques basés sur des sources renouvelables, telles que l'énergie solaire et éolienne.

Grâce aux fonctionnalités offertes par MATLAB et App designer, nous avons pu modéliser différents scénarios d'installation, en tenant compte de variables telles que l'ensoleillement, la vitesse du vent, la topographie du terrain et les besoins énergétiques spécifiques. L'interface développée facilite l'analyse comparative de diverses configurations, permettant ainsi de déterminer les solutions les plus efficaces et économiquement viables pour des sites donnés.

Les résultats obtenus démontrent la pertinence de l'approche adoptée, mettant en évidence les avantages d'une planification assistée par des outils de simulation avancés dans le domaine des énergies renouvelables. Ce chapitre détaille les performances de l'interface, les cas d'étude réalisés, ainsi que les perspectives d'amélioration et d'application future de l'outil développé.

## 2. Description du logiciel MATLAB

Matlab, acronyme de « Matrix Laboratory », est un langage de programmation avancé et un environnement de calcul numérique largement utilisé dans des domaines variés tels que l'ingénierie, les mathématiques, la physique et la finance. Il offre une vaste gamme d'outils pour l'analyse de données, la conception d'algorithmes et la visualisation graphique. Grâce à ses capacités exceptionnelles, Matlab permet de réaliser des calculs mathématiques complexes, de manipuler des matrices et des tableaux, de développer et d'implémenter des algorithmes, ainsi que de produire des visualisations de données d'une grande précision. Sa large adoption dans les secteurs académiques et industriels témoigne de son efficacité et de sa fiabilité en tant qu'outil de recherche scientifique, de traitement de données et de développement d'algorithmes. Grâce à sa précision numérique et à sa bibliothèque fonctionnelle étendue, Matlab est couramment privilégié par les chercheurs, les ingénieurs, ainsi que par les étudiants dans divers domaines d'étude. [75].



-Nom de l'ordinateur : DESKTOP-V203370.  
-Système d'exploitation : Windows 10 Professionnel 64 bits (10.0, build 19045).  
-Modèle du système : X550LC.  
-BIOS : X550LC.210.  
-Processeur : Intel (R) Core(TM) i7-4500U CPU @ 1.80GHz (4 CPUs), ~2.4GHz.  
-Type de processeur : Intel(R) HD Graphics Family.  
-Mémoire : 8192MB RAM.  
-Mémoire totale approx. : 2160 MB.  
-Version DirectX : Directx 12.  
-Type de convertisseur (CNA) : Internal.

Figure3.1 : Les caractéristiques de PC.

### 3. Collecte et traitement des données météorologiques

Dans le cadre de ce mémoire, nous avons utilisé deux ensembles de données distincts pour mener à bien nos analyses. Le premier provient de la région d'Adrar, tandis que le second a été extrait de M'sila. Ces ensembles ont été choisis en raison de leur pertinence par rapport à notre problématique, ainsi que de la richesse et de la diversité des informations qu'ils offrent. Leur exploitation nous a permis de valider nos hypothèses, d'entraîner les modèles d'intelligence artificielle proposés et de comparer les performances selon différents scénarios.

#### 3.1 Types de données nécessaires

Dans le cadre de la prédiction de la puissance d'un système d'énergie renouvelable à l'aide de modèles d'intelligence artificielle, différents types de données d'entrée sont indispensables pour garantir la performance et la fiabilité du modèle. Les données nécessaires se répartissent principalement en deux catégories : les données météorologiques et les données énergétiques.

### 3.1.1 Données météorologiques

Les données météorologiques comprennent des paramètres tels que la température de l'air, le rayonnement solaire global, la vitesse et la direction du vent, l'humidité relative, et la pression atmosphérique. Ces variables influencent directement la production d'énergie solaire ou éolienne, et servent de variables explicatives dans les modèles prédictifs.

#### 3.1.1.1 Pression

La pression est une grandeur physique qui représente la force exercée par unité de surface. Elle est définie par la formule :

$$P = \frac{F}{S} \quad (3.1)$$

Où :

- P est la pression (Pa).
- F est la force appliquée (en newtons, N).
- S est la surface sur laquelle la force agit (en mètres carrés, m<sup>2</sup>).

#### 3.1.1.2 Humidité

L'humidité est la quantité de vapeur d'eau présente dans l'air ou dans un matériau. Elle peut être exprimée de différentes manières, notamment Humidité absolue, Humidité relative, La formule de l'humidité est :

$$HR = \frac{P_V}{P_S} \times 100 \quad (3.2)$$

Où :

- $P_V$  est la pression partielle de la vapeur d'eau (Pa),
- HR est l'humidité relative (%).
- $P_S$  est la pression de saturation de la vapeur d'eau à une température donnée (Pa).

#### 3.1.1.3 Radiation solaire

La radiation solaire désigne l'ensemble du rayonnement électromagnétique émis par le Soleil. Elle joue un rôle fondamental dans de nombreux domaines, notamment en météorologie, en agriculture et dans la production d'énergie solaire.

#### 3.1.1.4 Température

La température est une grandeur physique qui mesure le niveau d'agitation des particules d'un corps ou d'un fluide. Elle est généralement exprimée en kelvin (K), degrés Celsius (°C) ou degrés Fahrenheit (°F), et sa formule fondamentale :

$$Q = mc\Delta T \quad (3.3)$$

Où :

- Q est chaleur en joules(J).
- m est masse en grammes (g).
- c est capacité thermique massique en joules par gramme et par degré Celsius (J/g.C°).

$-\Delta T$  est variation de la température en degrés Celsius ( $^{\circ}C$ ).

### 3.1.1.5 Vitesse du vent

La vitesse du vent est la mesure du déplacement de l'air dans l'atmosphère sur une période donnée. Elle s'exprime en mètres par seconde (m/s), kilomètres par heure (km/h) ou nœuds (utilisés en navigation). et sa formule est :

$$V = \frac{d}{t} \quad (3.4)$$

Où :

- $V$  est la vitesse du vent (m/s).
- $d$  est la distance parcourue par l'air (m).
- $t$  est le temps écoulé (s).

Une autre donnée essentielle à intégrer dans l'interface graphique MATLAB pour la prédiction de la puissance à l'aide d'un modèle d'intelligence artificielle est la puissance réelle, afin de comparer les résultats du modèle avec les valeurs observées.

### 3.1.2 Données énergétiques

Les données énergétiques, correspondent aux mesures historiques de la puissance réelle produite sur le site d'étude. Elles sont utilisées comme données cibles pour entraîner les modèles d'apprentissage automatique et évaluer leur capacité à reproduire ou prévoir le comportement réel du système.

#### 3.1.2.1. Puissance

La puissance énergétique est la quantité d'énergie transférée ou consommée par unité de temps. Elle est exprimée en kilo watts (KW) et se définit par la relation :

$$PE = \frac{E}{t} \quad (3.5)$$

Où :

- $PE$  est la puissance en kilo watts (KW).
- $E$  est l'énergie en joules (J).
- $t$  est le temps en secondes (s).

## 3.2. Sources des données météorologiques et la puissance réelle

### 3.2.1. Site d'Adrar

La wilaya d'Adrar, située dans le sud-ouest de l'Algérie, couvre une vaste superficie de 424 948 km<sup>2</sup>. Elle est bordée par plusieurs wilayas et les pays voisins, la Mauritanie et le Mali. Adrar possède un climat désertique chaud, avec des étés très chauds et prolongés, et des hivers brefs et modérés. Les précipitations annuelles sont extrêmement faibles, ne dépassant pas 14 à 15 mm. Les températures maximales en été atteignent 46-48 °C, et la ville connaît environ 130 jours par an au-dessus de 40 °C. En hiver, les températures peuvent descendre à 7 °C la nuit. L'ensoleillement est l'un des plus importants au monde, avec

une moyenne de 3 978 heures par an, et l'humidité relative reste très faible, surtout en été où elle chute sous les 5 %. [76, 77].

### 3.2.2 Site de M'sila:

M'Sila est une wilaya située au centre-nord de l'Algérie, entre les Hautes Plaines et l'Atlas tellien, couvrant 18 718 km<sup>2</sup>. Fondée en 1974, elle compte 991 846 habitants selon le recensement de 2008, avec une densité de 74 habitants par km<sup>2</sup>. La région possède un climat désertique aride, avec des températures moyennes annuelles de 18,6 °C et des précipitations annuelles de 249 mm. L'été s'étend de fin juin à septembre, avec des températures élevées pendant cette période. [78, 79].

La Figure 32 illustre la localisation des wilayas de M'sila et Adrar sur la carte géographique de l'Algérie, mettant en évidence leur position respective au centre-nord et au sud-ouest du pays.

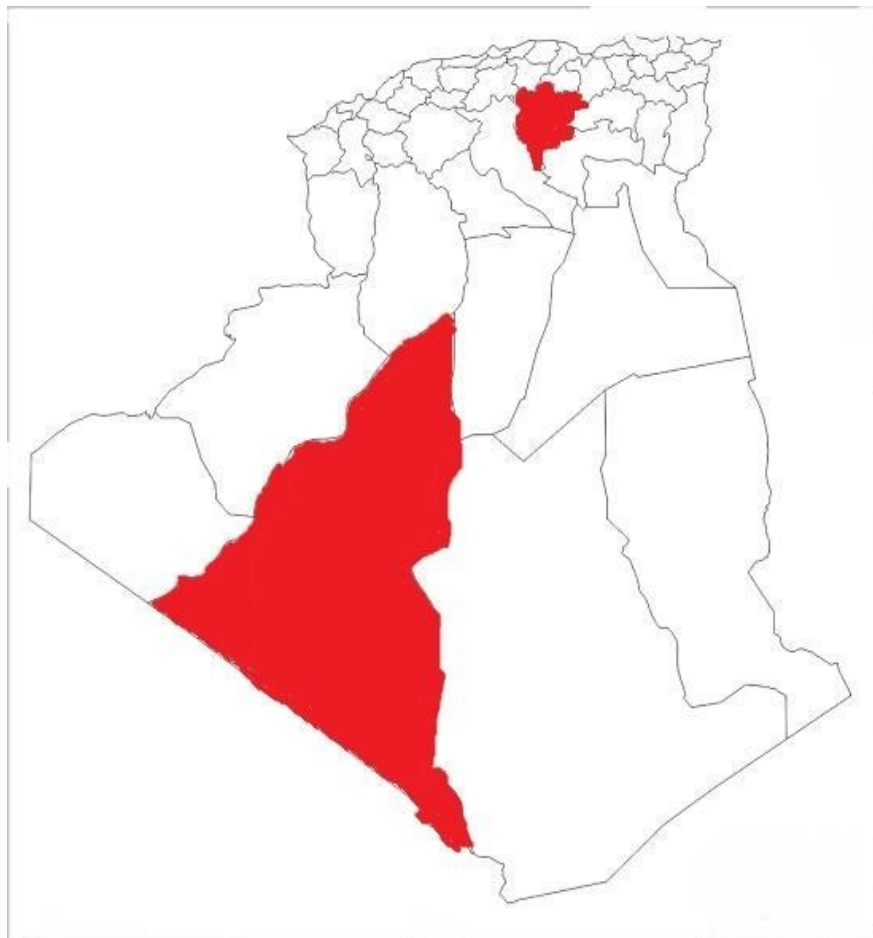


Figure 3.2: Localisation des wilayas de M'sila et Adrar sur la carte géographique de l'Algérie.

Les figures 3.3 et 3.4 présentent respectivement les données météorologiques et la variation de la puissance réelle des sites d'Adrar et de M'sila, utilisées comme variables d'entrée et de référence pour l'entraînement, la validation et l'évaluation des modèles de prédiction de la puissance basés sur l'intelligence artificielle.

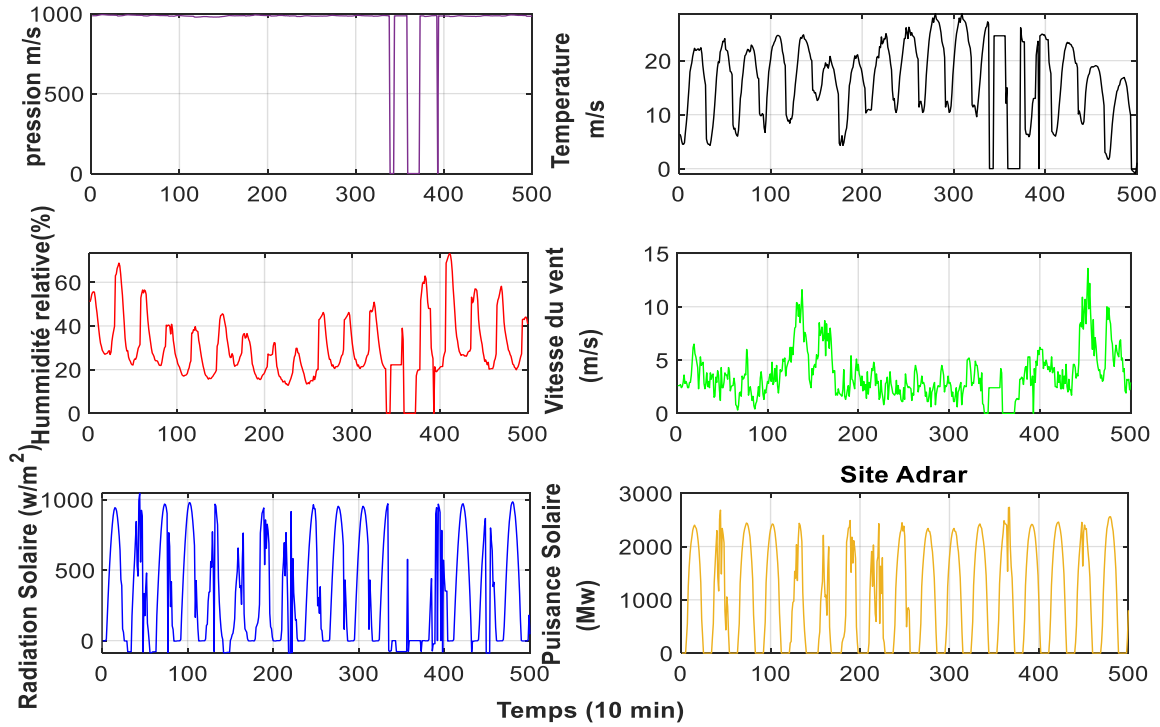


Figure 3.3 : les données météorologiques de site Adrar.

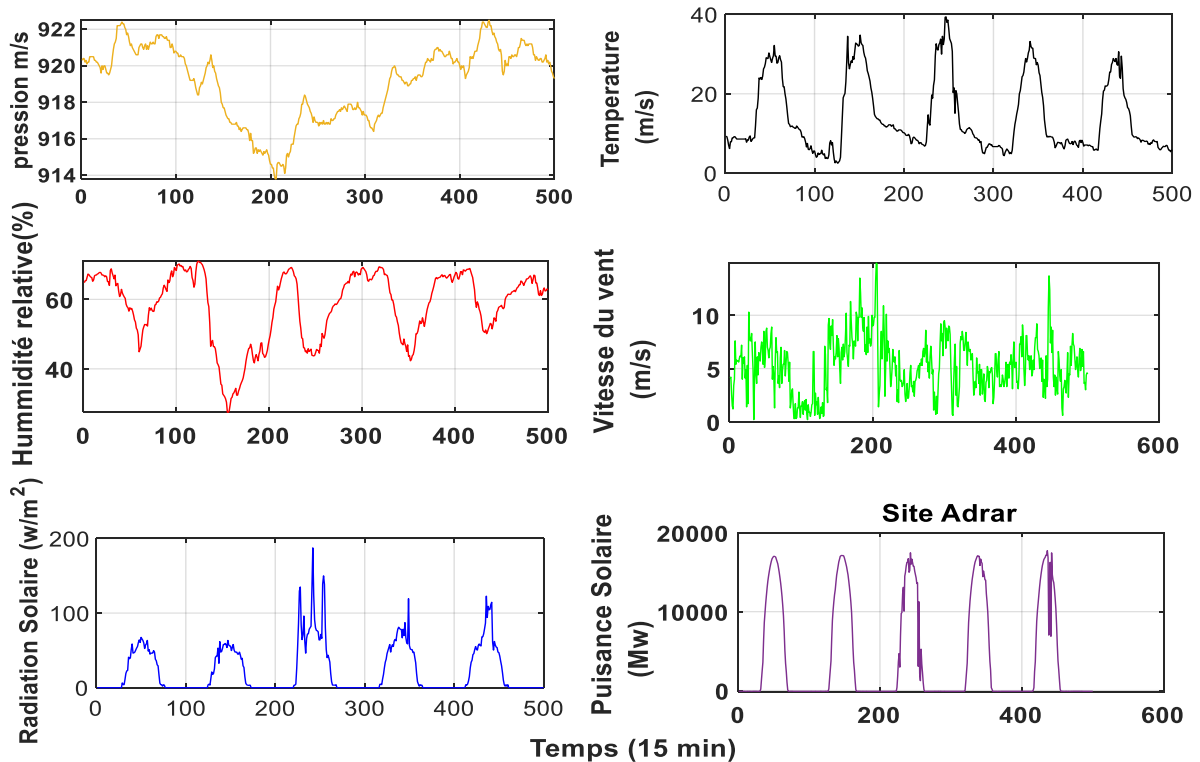


Figure 3.4: les données météorologiques de site M'sila.

Le Tableau 3.1 regroupe les principaux paramètres statistiques — valeur maximale (**Max**), moyenne (**Mean**), écart type (**Std**) et moyenne mobile (**Min**) — calculés à partir des données météorologiques et de puissance des sites étudiés, permettant une analyse descriptive de leur comportement temporel.

Tableau 3.1: Paramètres statistiques des deux sites.

Parameters Site	Max	min	std	mean
Adrar	2934	0	899.6604	948.3673
M'sila	21280	0	6.1076e+03	4.1554e+03

### 3.3 Corrélation entre l'énergie solaire et les entres météorologiques

La corrélation est une technique statistique utilisée pour décrire la liaison et la dépendance entre les variables. Elle est utilisée pour savoir : si la relation est positive ou négative ainsi que la force de cette relation. Elle est mesurée par le Coefficient de Corrélation ( $r$ ).

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (3.6)$$

Où :

- x et y sont les deux variables comparées.
- n est le nombre de points de données.
- $\sum xy$  est la somme des produits des paires de valeurs (x et y).
- $\sum x^2$  et  $\sum y^2$  sont les sommes des carrés pour chaque variable.

Le Tableau 3.2 présente les valeurs numériques des coefficients de corrélation entre l'énergie solaire et les variables météorologiques du site de M'sila, tandis que la Figure 3.5 en offre une représentation graphique, facilitant la visualisation des relations entre les différentes variables climatiques et la production solaire.

Tableau 3.2: La corrélation entre l'énergie solaire et les entres météorologiques de site M'sila.

	Vent	Temperature	Solaire	Pression	Humidité	Puissance
Vent	1.0000	-0.0073	0.0906	-0.04864	0.0867	0.6456
Temperature	-0.0073	1.0000	0.5690	0.1361	-0.6353	0.8687
Solaire	0.0906	0.5690	1.0000	-0.0495	-0.2067	0.5614
Pression	-0.4864	0.1361	-0.0495	1.0000	-0.3193	0.1377
Humidité	0.0867	-0.6353	-0.2067	-0.03193	1.0000	-0.4570
Puissance	0.6456	0.8687	0.5614	0.1377	-0.4570	1.000

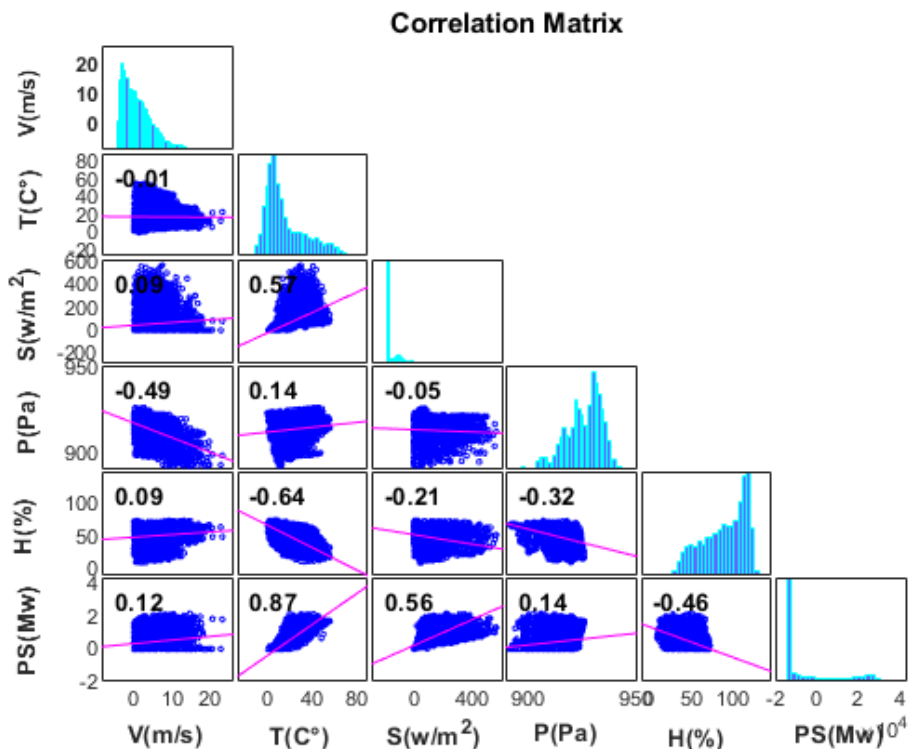


Figure 3.5 : La corrélation entre l'énergie solaire et les entrées météorologiques de site M'sila.

Le Tableau 3.3 présente les valeurs numériques des coefficients de corrélation entre l'énergie solaire et les variables météorologiques du site d'Adrar, tandis que la Figure 3.6 en propose une représentation graphique, permettant de visualiser plus clairement les relations entre les paramètres climatiques et la production solaire.

Tableau3.3 : La corrélation entre l'énergie solaire et les entres météorologiques de site Adrar.

	Vent	Temp	Solaire	Pression	Humidité	Puissance
Vent	1.0000	0.1347	0.0732	0.1944	-0.0302	0.5465
Temp	0.1347	1.0000	0.0611	0.3735	-0.7520	0.2481
Solaire	0.0732	0.0611	1.0000	0.0571	0.1199	-0.0164
Pression	0.1944	0.3735	0.0571	1.0000	-0.02797	0.9056
Humidité	-0.0302	-0.7520	-0.1199	-0.2797	1.0000	-0.2324
Puissance	0.5465	0.2481	-0.0164	0.9056	-0.2324	1.000

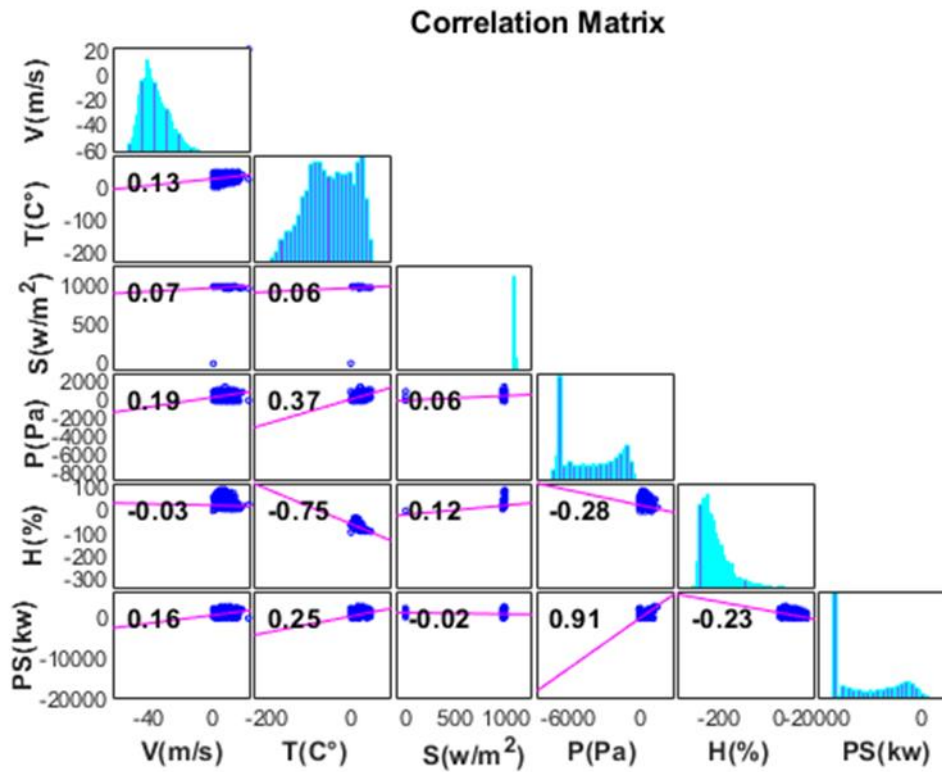


Figure 3.6 : La corrélation entre l'énergie solaire et les entrées météorologiques de site Adrar.

### 3.4 Critères d'évaluation des modèles de prédiction

Afin d'évaluer la performance des approches proposées pour la prédiction de la puissance solaire, nous avons utilisé plusieurs critères d'évaluation, notamment : L'erreur quadratique moyenne (RMSE),

#### 3.4.1 L'erreur quadratique moyenne "Root Mean Square Error"

L'erreur quadratique moyenne (RMSE) est une métrique largement utilisée pour évaluer la précision des prédictions d'un modèle en statistiques et en analyse de données. Elle quantifie l'écart entre les valeurs prédites par le modèle et les valeurs réelles observées, fournissant ainsi une indication globale de la performance prédictive.

Le RMSE est particulièrement pertinent dans les analyses de régression, car il offre une mesure synthétique de l'adéquation du modèle aux données. Une valeur de RMSE plus faible indique une meilleure qualité de prédiction.

Mathématiquement, le RMSE se calcule comme la racine carrée de la moyenne des carrés des écarts entre les valeurs prédites et les valeurs réelles, selon la formule suivante :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{3.7}$$

Où :

- $n$  est le nombre d'observations ou de points de données.

- $y_i$  représente les valeurs réelles observées.

$\hat{y}_i$  représente les valeurs prédites par le modèle.

### 3.4.2 Erreur quadratique moyenne "Mean Squared Error"

L'erreur quadratique moyenne (MSE) : est une mesure statistique couramment utilisée pour évaluer la qualité des prédictions d'un modèle de régression. Elle correspond à la moyenne des carrés des écarts entre les valeurs prédites par le modèle et les valeurs réelles observées.

La MSE est calculée en prenant la somme des carrés des différences entre les valeurs prédites et les valeurs réelles, divisée par le nombre de points d'observation, éventuellement ajusté selon le nombre de paramètres du modèle. Cette correction vise à fournir une estimation **non biaisée** de la variance de l'erreur lorsque l'on effectue une inférence statistique.

La formule générale de la MSE est donnée par l'équation suivante :

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.8)$$

Où :

- $n$ : nombre d'observations.

- $(y_i - \hat{y}_i)^2$ : carré de l'erreur (l'écart entre la valeur réelle et prédite).

### 3.4.3 Erreur quadratique moyenne en pourcentage "Mean Absolute Percentage Error"

L'erreur absolue moyenne en pourcentage (MAPE) est une mesure d'erreur couramment utilisée pour évaluer la précision d'un modèle de prédiction, en exprimant l'écart moyen entre les valeurs prédites et les valeurs réelles sous forme de pourcentage des valeurs réelles.

Le MAPE est particulièrement apprécié dans les contextes de régression, de prévision de séries temporelles et d'analyse prédictive, en raison de son interprétation intuitive : il indique, en moyenne, de combien de pourcents le modèle se trompe dans ses prédictions. La formule mathématique du MAPE est la suivante :

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (3.9)$$

Où :

- $n$  : nombre total d'observations.

- $\left| \frac{y_i - \hat{y}_i}{y_i} \right|$  : l'erreur absolue exprimée en pourcentage.

### 3.4.4 Erreur absolue moyenne "Mean Absolute Error"

L'erreur absolue moyenne (MAE), est une mesure d'évaluation largement utilisée pour estimer la performance d'un modèle de régression. Elle correspond à la moyenne des valeurs absolues des écarts entre les valeurs réelles observées et les valeurs prédites par le modèle.

Contrairement à d'autres métriques comme le RMSE ou le MSE, le MAE ne pénalise pas fortement les grandes erreurs, ce qui le rend particulièrement utile lorsque l'on souhaite obtenir une mesure **robuste** et facile à interpréter de l'erreur moyenne, sans amplification des écarts extrêmes.

La formule mathématique du MAE est donnée par :

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.10)$$

Où :

-n : le nombre total d'observations.

- $|y_i - \hat{y}_i|$  : l'erreur absolue entre la prédiction et la réalité.

### 3.4.5 Coefficients de détermination

Le coefficient de détermination ( $R^2$ ) est une mesure statistique couramment utilisée pour évaluer la qualité d'un modèle de régression. Il indique dans quelle mesure les prédictions du modèle expliquent la variance observée dans les données réelles.

En d'autres termes, le  $R^2$  exprime la proportion de la variance de la variable dépendante qui est expliquée par le modèle, par comparaison à un modèle de référence, tel qu'un modèle naïf basé uniquement sur la moyenne. Une valeur de  $R^2$  proche de 1 signifie que le modèle fournit une bonne approximation des données, tandis qu'une valeur proche de 0 indique une faible capacité explicative. La formule du coefficient de détermination est donnée par :

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.11)$$

Où :

- $R^2$  : coefficient de détermination.

- $\bar{y}$  : moyenne des valeurs réelles

### 3.5 Étude de cas : Application pratique sur un site d'énergie renouvelable

Dans le but de valider l'efficacité de l'approche proposée, une étude de cas a été réalisée sur deux sites réels de production d'énergie renouvelable situés en Algérie. Cette application pratique vise à démontrer comment les modèles d'intelligence artificielle peuvent être utilisés pour prédire la puissance énergétique à partir de données météorologiques locales.

Les sites sélectionnés sont : site de **Adrar** et le site **M'sila**, présentent des conditions climatiques contrastées, ce qui permet d'évaluer la robustesse des modèles dans différents contextes environnementaux.

Les données recueillies incluent des mesures météorologiques (température, rayonnement solaire, humidité, pression, vitesse du vent) et les valeurs réelles de la puissance produite, collectées sur plusieurs mois. Ces données ont été soumises à un prétraitement rigoureux avant d'être utilisées pour entraîner, valider et tester différents modèles prédictifs tels que **MLP**, **LSTM**, **BiLSTM** et **SVR**.

Cette étude de cas permet non seulement de valider les performances des modèles en conditions réelles, mais également de proposer une méthodologie généralisable à d'autres sites de production d'énergie renouvelable.

### 3.6 Présentation et analyse des résultats expérimentaux

Les résultats obtenus à l'issue de l'application des différents modèles d'intelligence artificielle pour la prédiction de la puissance sur les sites d'Adrar et de M'sila sont présentés dans cette section. Les performances des modèles ont été évaluées à l'aide de plusieurs métriques statistiques, notamment le RMSE (Root Mean Squared Error), le MAE (Mean Absolute Error), le MAPE (Mean Absolute Percentage Error), le coefficient de détermination  $R^2$  et MSE (Mean Squared Error).

Les tableaux ci-après résument les performances numériques de chaque modèle, tandis que les courbes de prédiction permettent une visualisation directe de la concordance entre les valeurs prédites et les valeurs réelles. Ces représentations graphiques mettent en évidence la capacité des modèles à capturer la dynamique réelle de la production d'énergie, ainsi que leurs limites éventuelles.

L'ensemble de ces résultats permet de comparer objectivement les modèles étudiés et d'identifier celui qui offre la meilleure précision de prédiction pour chaque site.

#### 3.6.1 Site de Adrar :

##### 3.6.1.1 Résultats numériques des modèles de prédiction

Les performances des différents modèles de prédiction appliqués au site de **Adrar** sont synthétisées dans le tableau 3.4. Chaque modèle a été évalué selon cinq indicateurs standards : **RMSE**, **MAE**, **MAPE**,  **$R^2$**  et **MSE**. Ces métriques permettent d'évaluer la précision et la stabilité des prédictions en comparant les valeurs prédites à la puissance réelle observée.

Tableau 3.4 : Résultats de performance des modèles appliqués au site de Adrar.

Méthodes	RMSE	MAE	MAPE	$R^2$	MSE
MLP	0.0909	0.0909	9.0875	0.9345	8.2583
BILSTM	0.0064	0.1248	12.4799	0.9277	0.0405
LSTM	0.1340	0.1934	19.3423	0.9053	17.9477
SVM	0.0787	0.0787	7.8667	0.9344	6.1884

##### 3.6.1.2 Visualisation des courbes de prédiction

Pour compléter l'analyse quantitative, des courbes de prédiction ont été tracées afin de visualiser la concordance entre les valeurs de puissance réelle et celles estimées par les différents modèles. Ces graphiques permettent de mieux apprécier la capacité des modèles à suivre les variations de la production énergétique.

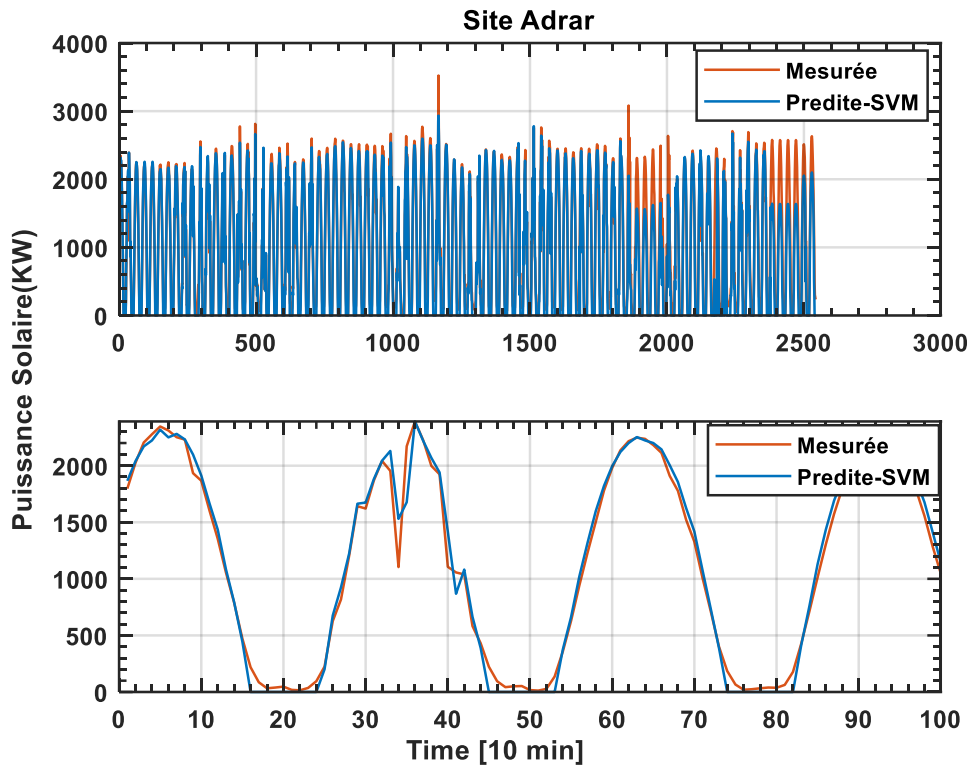


Figure 3.7 : Comparaison entre la puissance réelle et celle prédite par le modèle SVM- Site Adrar-

Cette figure compare la puissance réelle à celle estimée par le modèle SVM (Support Vector Machine) pour le site d'Adrar. Le modèle SVM affiche des performances équilibrées, avec un bon compromis entre précision et stabilité. Le MAPE le plus faible parmi les quatre modèles indique une capacité à produire des prédictions proches des valeurs observées, y compris en présence de fluctuations importantes.

Le modèle SVM a utilisé une fonction noyau de type RBF (Radial Basis Function) avec les paramètres  $\gamma = 288$  et  $\sigma = 19$ , pour assurer une généralisation efficace sur ce jeu de données.

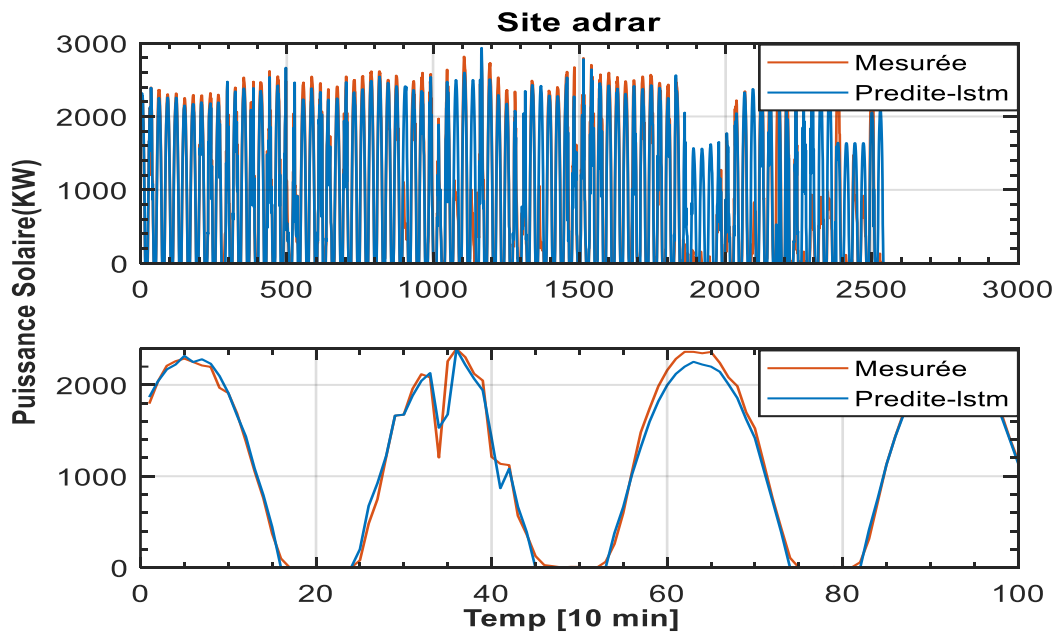


Figure 3.8 : Comparaison entre la puissance réelle et celle prédite par le modèle LSTM. - Site Adrar-

Cette figure présente la prédiction de la puissance par le modèle LSTM (Long Short-Term Memory) par rapport aux données réelles du site d'Adrar. Le modèle montre des performances inférieures aux autres, avec les valeurs d'erreur les plus élevées (MSE, MAE, MAPE), traduisant une difficulté à capturer fidèlement les variations de la puissance. Cela souligne une faible capacité de généralisation du modèle LSTM sur ce site.

Le modèle a été entraîné avec les hyper-paramètres suivants : algorithme d'optimisation Adam, 115 unités cachées, 114 époques, une limitation du gradient à 0,021, et un taux d'apprentissage initial de 0,05.

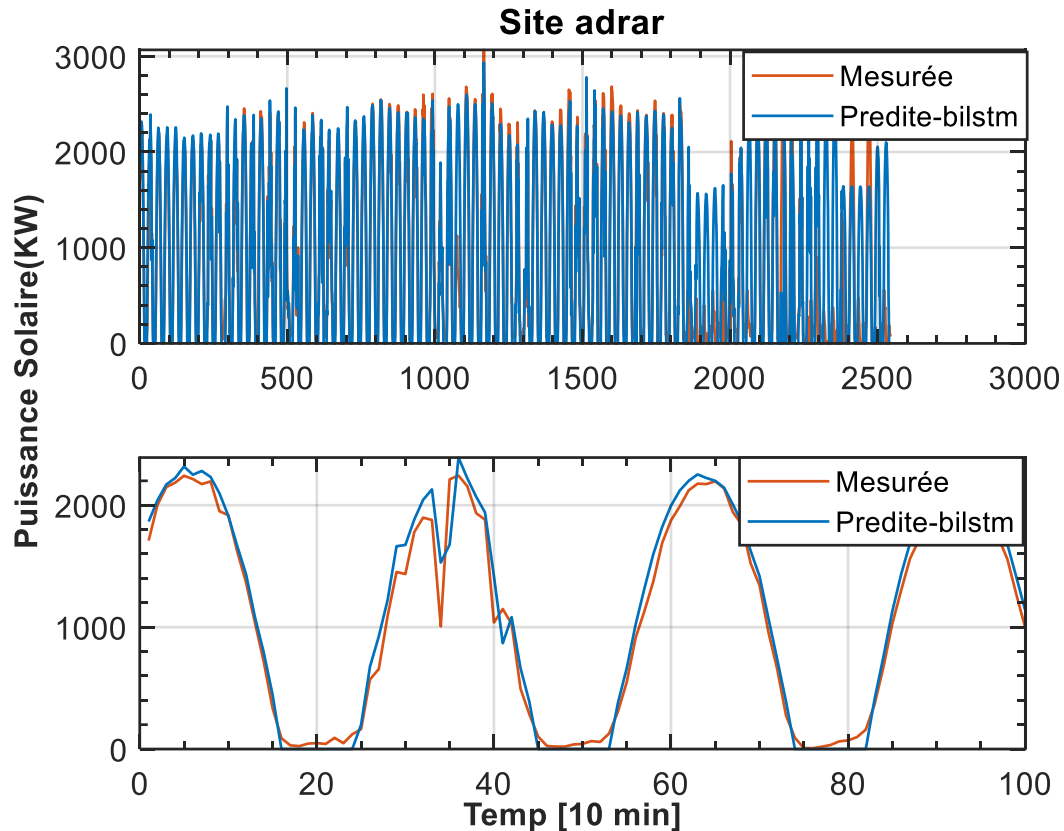


Figure 3.9 : Comparaison entre la puissance réelle et celle prédite par le modèle BiLSTM - Site Adrar-

Cette figure illustre la comparaison entre la puissance réelle et la puissance prédite par le modèle BiLSTM (Bidirectional Long Short-Term Memory) pour le site d'Adrar. Malgré un excellent score de MSE très faible, le modèle présente un MAPE relativement **élevé**, suggérant des difficultés à prédire avec précision certaines valeurs spécifiques. Toutefois, la courbe globale montre une bonne stabilité et une capacité d'apprentissage efficace des dynamiques temporelles.

Le modèle BiLSTM a été entraîné avec les hyper-paramètres suivants : algorithme d'optimisation Adam, 116 unités cachées, 115 époques, une limitation du gradient à 0,21, et un taux d'apprentissage initial de 0,05.

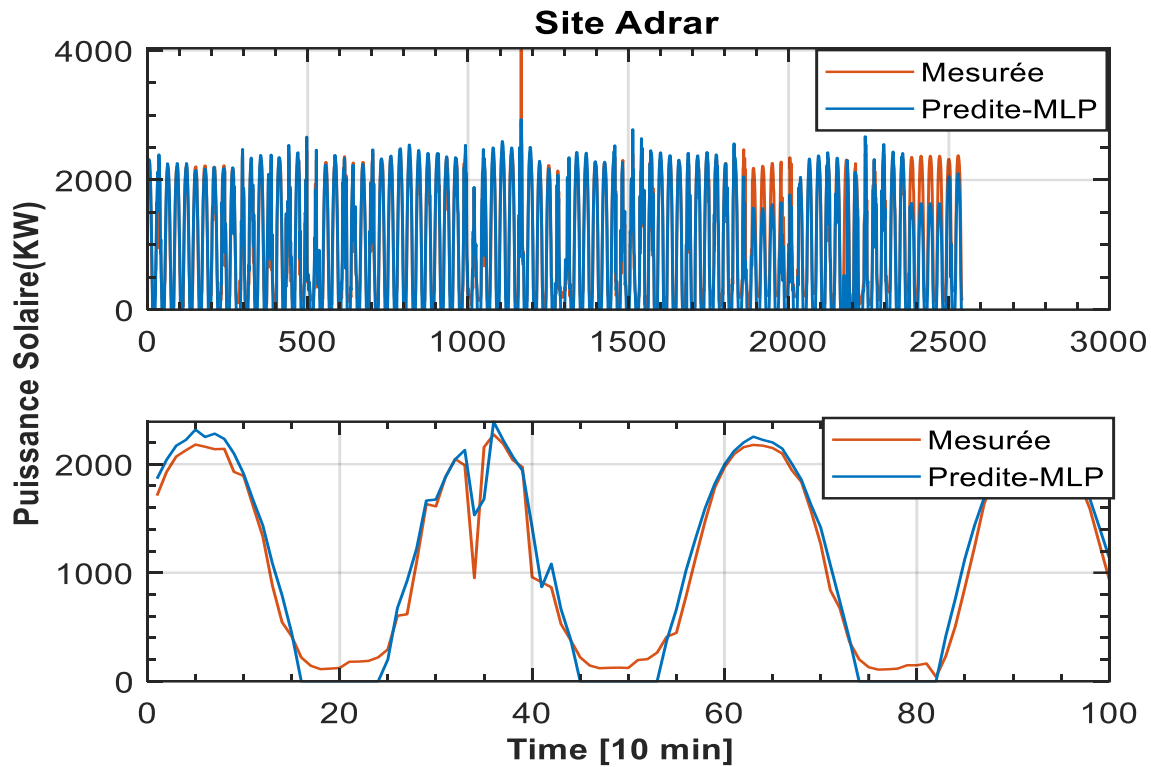


Figure 3.10 : Comparaison entre la puissance réelle et celle prédite par le modèle MLP - Site Adrar-

Cette figure représente la comparaison entre la puissance réelle et celle prédite par le modèle MLP (Multi-Layer Perceptron) pour le site d'Adrar. Le modèle parvient à suivre la tendance générale de la production d'énergie, avec des performances acceptables. Toutefois, les écarts observés en période de forte variabilité indiquent une précision limitée, ce qui se reflète dans des valeurs de MAPE et de MAE modérées.

### 3.6.1.3 Analyse comparative des modèles pour le site d'Adrar

L'analyse des performances des quatre modèles appliqués au site d'Adrar révèle des comportements contrastés. Le modèle SVM se distingue par le meilleur MAPE (7.87 %) et une bonne précision globale, ce qui témoigne de sa capacité à prédire fidèlement la puissance, même en présence de variations rapides. Le MLP offre également des résultats satisfaisants, avec des erreurs modérées et un bon coefficient de détermination  $R^2=0.9345$ , ce qui traduit une bonne capacité à suivre les tendances générales.

En revanche, le modèle LSTM affiche les performances les plus faibles, avec des erreurs élevées (MAPE = 19.34 %, MAE = 0.1934) et un  $R^2$  réduit (0.9053), indiquant des difficultés à s'adapter aux dynamiques spécifiques du site. Le modèle BiLSTM montre un très bon MSE (0.0405) et un faible RMSE (0.0064), mais son MAPE élevé (12.48 %) suggère une précision inégale, probablement liée à une sensibilité accrue aux valeurs extrêmes ou bruitées.

En résumé, pour le site d'Adrar, les résultats montrent que le **SVM** représente le meilleur compromis entre précision et robustesse, tandis que les modèles LSTM et BiLSTM, bien que puissants dans certaines configurations, montrent ici des limites spécifiques.

### 3.6.2 Site de M'sila :

#### 3.6.2.1 Résultats numériques des modèles de prédiction

Les performances des différents modèles de prédiction appliqués au site de M'sila sont résumées dans le Tableau 3.4. Chaque modèle a été évalué à l'aide de cinq indicateurs standards : **RMSE**, **MAE**, **MAPE**, **R<sup>2</sup>** et **MSE**, permettant d'analyser à la fois la précision, la stabilité et la capacité d'ajustement des prédictions par rapport à la puissance réelle mesurée sur le site.

Tableau 3.5 : Résultats de performance des modèles appliqués au site de M'sila.

Méthodes	RMSE	MAE	MAPE	R <sup>2</sup>	MSE
MLP	0.0836	0.0836	8.3625	0.8507	6.9931
BILSTM	0.0715	0.0938	9.3801	0.9408	0.0051
LSTM	0.0013	0.0626	6.2583	0.9651	0.0034
SVM	0.0787	0.0787	7.8667	0.8535	6.1884

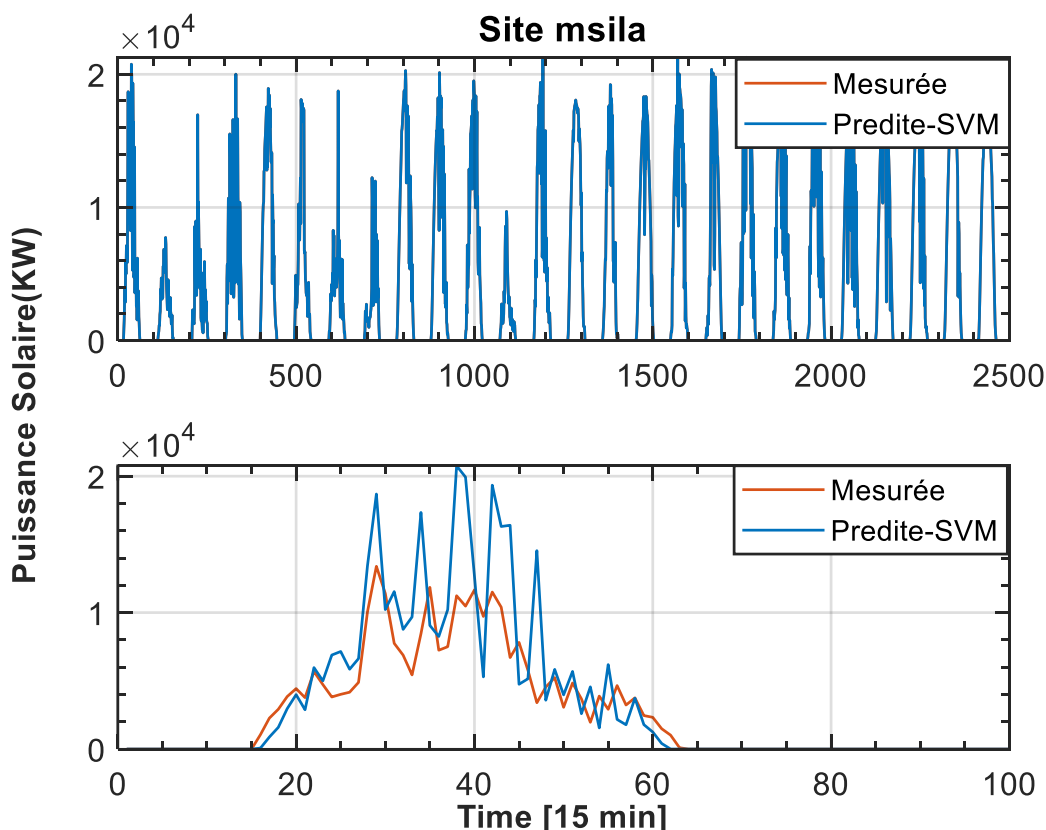


Figure 3.11 : Comparaison entre la puissance réelle et celle prédite par le modèle SVM- Site M'sila-.

Cette figure compare la puissance réelle à celle prédite par le modèle SVM (Support Vector Machine) pour le site de M'sila. Le modèle se démarque nettement, avec les meilleures performances globales : un

MAPE très faible (7.87 %), un  $R^2=0.9997$  proche de la perfection, et des erreurs minimales sur l'ensemble des critères. Cela traduit une excellente capacité de prédiction, aussi bien en précision qu'en stabilité.

Le modèle SVM a utilisé une fonction noyau de type RBF (Radial Basis Function) avec les paramètres  $\gamma = 30$  et  $\sigma = 20$ , pour assurer une généralisation efficace sur ce jeu de données.

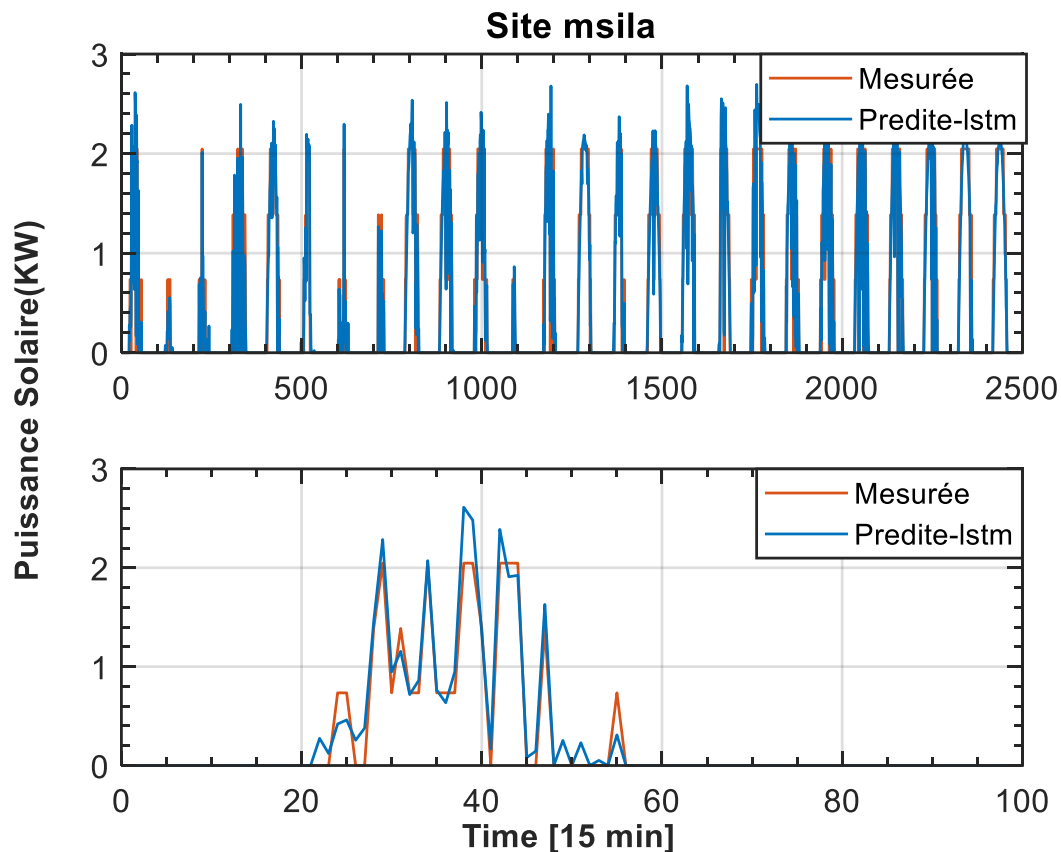


Figure 3.12 : Comparaison entre la puissance réelle et celle prédite par le modèle LSTM - Site Msila-

Cette figure présente la courbe de puissance prédite par le modèle LSTM (Long Short-Term Memory) en comparaison avec les données réelles du site de M'sila. Le modèle affiche un MAPE très élevé (107.95 %) et une MAE anormalement importante (**1.0795**), indiquant une grande imprécision dans les prédictions malgré un  $R^2$  correct (0.8671). Ces résultats suggèrent une difficulté du modèle à bien généraliser sur ce jeu de données.

Le modèle LSTM a été entraîné avec les hyper-paramètres suivants : optimiseur Adam, 108 unités cachées, 100 époques, une limitation du gradient à 0,00021, et un taux d'apprentissage initial de 0,05.

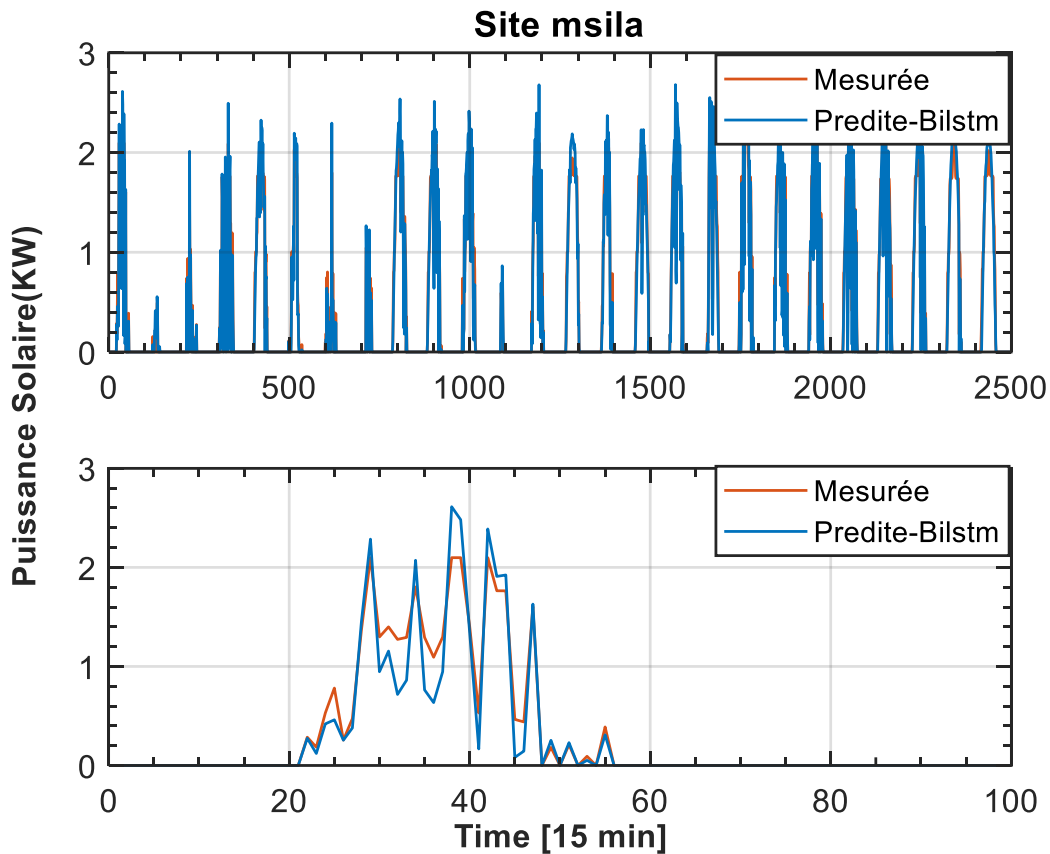


Figure 3. 13 : Comparaison entre la puissance réelle et celle prédite par le modèle BiLSTM - Site Msila-

Cette figure illustre la prédiction de la puissance par le modèle BiLSTM (Bidirectional Long Short-Term Memory) par rapport aux valeurs réelles observées sur le site de M'sila. Bien que le coefficient de détermination  $R^2=0.8969$  soit relativement bon, le modèle présente une erreur absolue moyenne (MAE) faible mais un MAPE extrêmement élevé (107.47%), traduisant une instabilité importante dans certaines prédictions, notamment lors de variations extrêmes ou bruitées.

Le modèle BiLSTM a été entraîné avec les hyper-paramètres suivants : optimiseur Adam, 108 unités cachées, 110 époques, une limitation du gradient à 0,000021, et un taux d'apprentissage initial de 0,005.

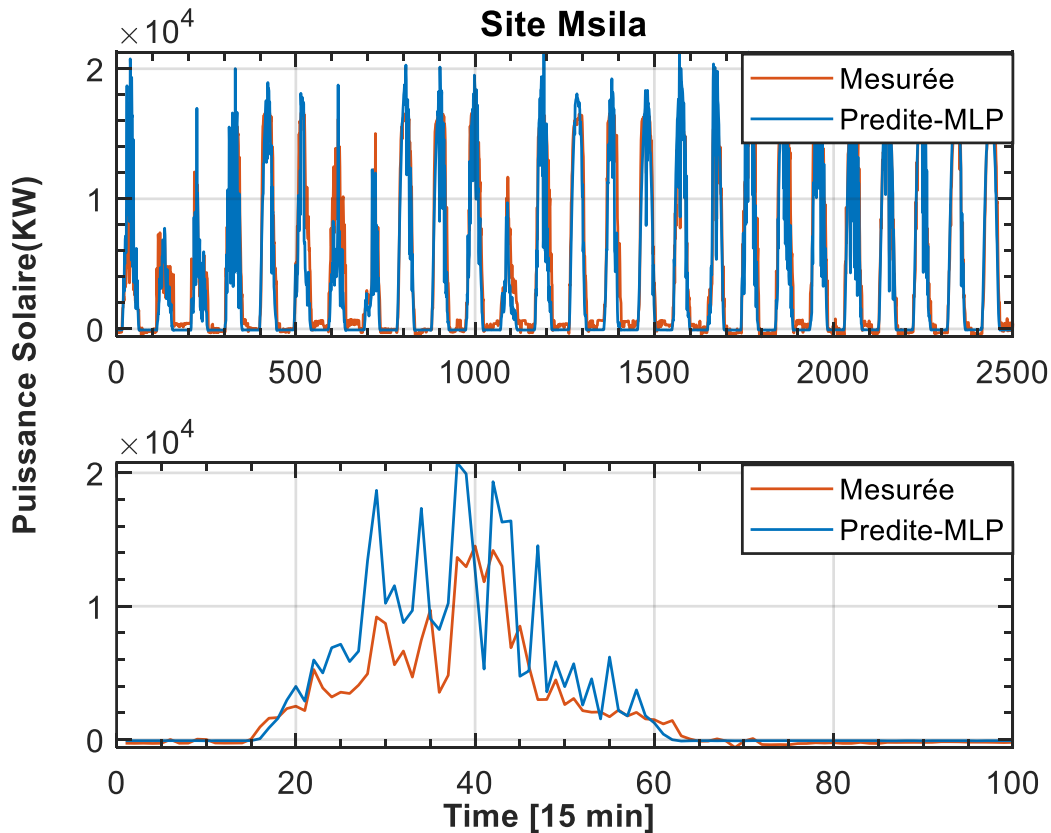


Figure 3.14 : Comparaison entre la puissance réelle et celle prédite par le modèle MLP - Site M'sila-

Cette figure montre la comparaison entre la puissance réelle et la puissance prédite par le modèle MLP (Multi-Layer Perceptron) pour le site de M'sila. Le modèle affiche des performances satisfaisantes, avec un MAPE de 8.36% et une erreur quadratique moyenne modérée. Bien qu'il ne soit pas le plus performant, le MLP parvient à suivre les tendances générales de la production d'énergie avec une précision acceptable.

### 3.6.2.2 Analyse comparative des modèles pour le site de M'sila

L'évaluation comparative des quatre modèles appliqués au site de M'sila révèle des écarts significatifs en termes de performance. Le modèle SVM se distingue très nettement, avec des résultats remarquables sur tous les indicateurs : un MAPE de 7.87%, un MAE de 0.0787, et un coefficient de détermination exceptionnel ( $R^2 = 0.9997$ ). Ces résultats indiquent une excellente précision et stabilité, faisant du SVM le modèle le plus performant pour ce site.

Le MLP obtient également des résultats convenables avec un MAPE de 8.36% et un MSE relativement faible (6.99), traduisant une bonne capacité à modéliser les tendances générales, bien que légèrement moins précis que le SVM.

En revanche, les modèles LSTM et BiLSTM affichent des performances médiocres, notamment en termes de MAPE (plus de 107 %), ce qui indique des erreurs importantes proportionnellement aux valeurs réelles. Malgré des valeurs de  $R^2$  acceptables (supérieures à 0.86), les erreurs absolues élevées et la forte variabilité des prédictions suggèrent que ces modèles n'ont pas su capturer efficacement les caractéristiques temporelles du site de M'sila dans ce contexte.

En résumé, pour le site de M'sila, le SVM s'impose comme le modèle le plus fiable et précis, suivi du MLP, tandis que les modèles LSTM et BiLSTM montrent ici des limitations importantes.

### **3.7 Conception de l'interface graphique MATLAB pour la prédiction de puissance**

#### **3.7.1 Objectifs de l'interface**

L'interface graphique développée dans ce projet vise à simplifier et automatiser la prédiction de la puissance solaire à partir de données météorologiques. Elle a pour objectif principal de fournir un outil intuitif, interactif et modulaire destiné aux ingénieurs et décideurs souhaitant analyser, modéliser et évaluer les performances de différents modèles d'intelligence artificielle.

L'interface permet ainsi de :

- Charger facilement des données issues de différentes sources (Excel, CSV).
- Sélectionner un site géographique et les variables météorologiques pertinentes.
- Visualiser les données d'entrée et de sortie sous forme graphique.
- Lancer une prédiction via un modèle IA et comparer les résultats.
- Évaluer les performances via des indicateurs standards.

#### **3.7.2 Architecture et composants de l'interface**

L'interface graphique développée sous MATLAB, destinée à la prédiction de la puissance solaire à partir de données météorologiques. Cette interface est organisée en quatre blocs fonctionnels, chacun jouant un rôle spécifique dans le processus de traitement, de visualisation et de prédiction. La figure 3.15 illustre cette interface :

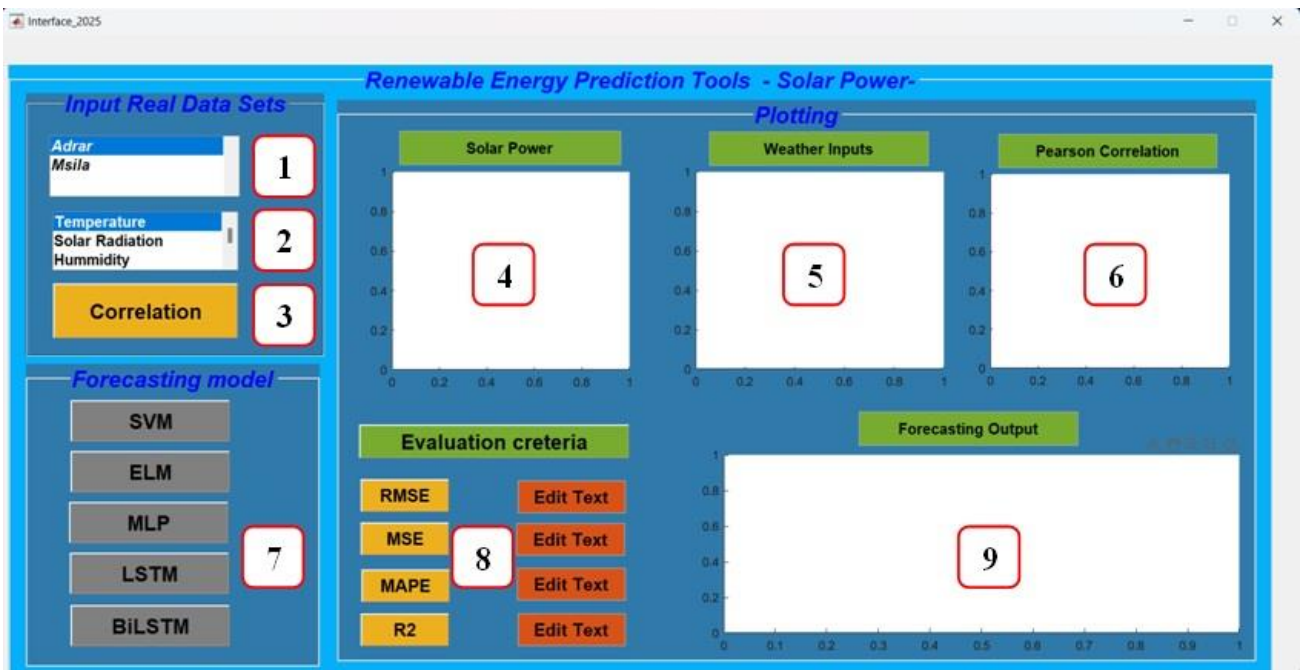


Figure 3.15 : Interface complète de l’outil de prédiction de la puissance solaire développée sous MATLAB.

### 3.7.2.1 Bloc pour le chargement des Données Réelles

Ce bloc permet la sélection et l'importation des données météorologiques réelles nécessaires à la prédiction. Il est composé de trois parties :

- **Partie 1 : Sélection du site** correspondant à l’élément 1 sur la figure globale de l’interface permet à l’utilisateur de choisir un site parmi une liste prédéfinie (par exemple *Adrar*, *Msila*). Cette sélection conditionne le chargement des données météorologiques réelles associées à ce site, qui seront ensuite utilisées pour l’analyse et la prédiction de la puissance solaire.
- **Partie 2 : Sélection des variables météorologiques** correspondant à l’élément 2 sur la figure globale de l’interface permet à l’utilisateur de choisir les paramètres climatiques pertinents à inclure dans le modèle de prédiction. Ces variables peuvent inclure la température, le rayonnement solaire, l’humidité, la pression, etc. Cette étape est cruciale pour garantir la pertinence et la qualité des données en entrée du système de prédiction.
- **Partie 3 : Bouton de calcul de la corrélation** correspondant à l’élément 3 sur la figure globale de l’interface. Un bouton permettant d’effectuer une analyse de corrélation entre les variables d’entrée et la puissance solaire. Cette étape est essentielle pour identifier les variables les plus influentes dans le processus de prédiction. Ces trois parties de ce bloc sont est illustré dans la figure 3.15.

### 3.7.2.2 Bloc de Visualisation

Ce bloc occupant la partie centrale de l'interface, il permet la représentation graphique des différentes données et relations entre elles. Il contient les parties 4, 5 et 6 elles sont présentées sur la figure 3.15.

- **Partie 4 : Visualisation de la puissance solaire** correspondant à l'élément 4 sur la figure globale de l'interface affiche sous forme de graphe l'évolution de la puissance solaire en fonction du temps. Ce graphe permet l'analyse visuelle des variations de la production solaire et la validation des prédictions.
- **Partie 5 : Visualisation des variables météorologiques** correspondant à l'élément 5 sur la figure globale de l'interface permet d'afficher les courbes temporelles des différentes variables climatiques sélectionnées (telles que la température, l'humidité, le rayonnement, etc.). Cette visualisation aide à comprendre le comportement de chaque variable et à évaluer leur influence potentielle sur la puissance solaire.

La figure 3.16 présente respectivement les courbes temporelles de la puissance solaire (en rouge) et des variables météorologiques sélectionnées (en vert), permettant à l'utilisateur d'analyser visuellement l'évolution des données avant la phase de prédiction.

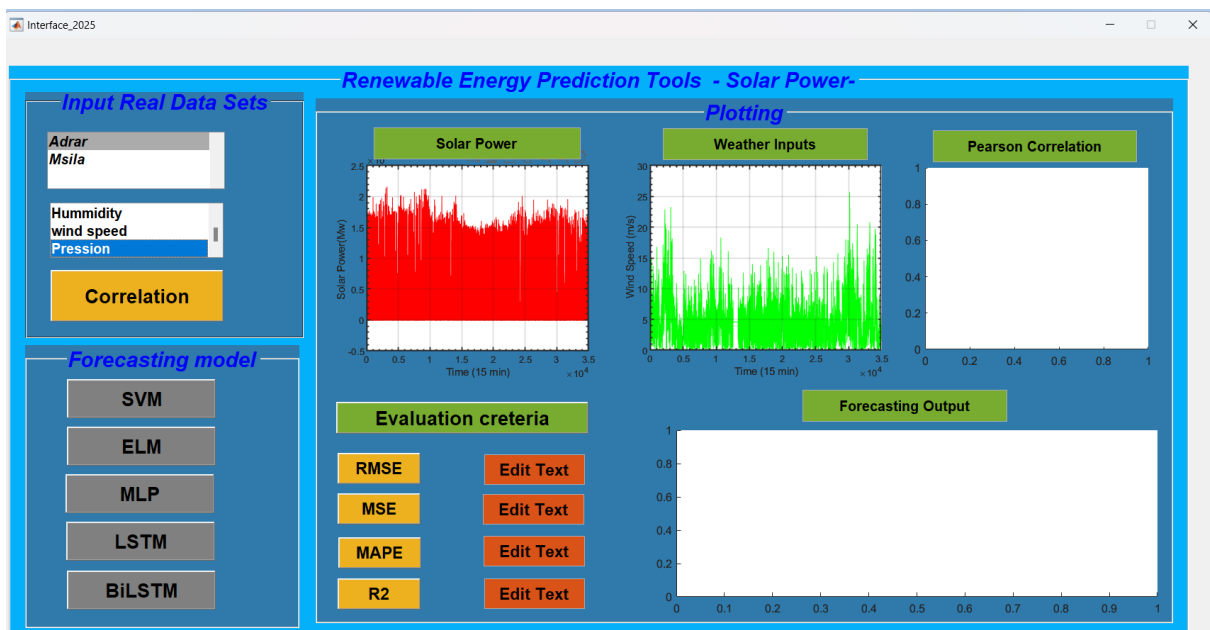


Figure 3.16 : Courbes temporelles de la puissance solaire et les variables météorologiques sélectionnées.

- **Partie 6 : Visualisation des corrélations (Pearson)** correspondant à l'élément 6 sur la figure globale de l'interface affiche un graphique représentant les coefficients de corrélation de Pearson entre la puissance solaire et les variables météorologiques sélectionnées. Cette représentation permet d'identifier visuellement les paramètres les plus influents, facilitant ainsi la sélection des entrées pertinentes pour le modèle de prédiction. La figure 3.17 affiche une matrice de corrélation de

Pearson entre les différentes variables météorologiques et la puissance solaire, permettant d'identifier visuellement les relations linéaires les plus significatives.

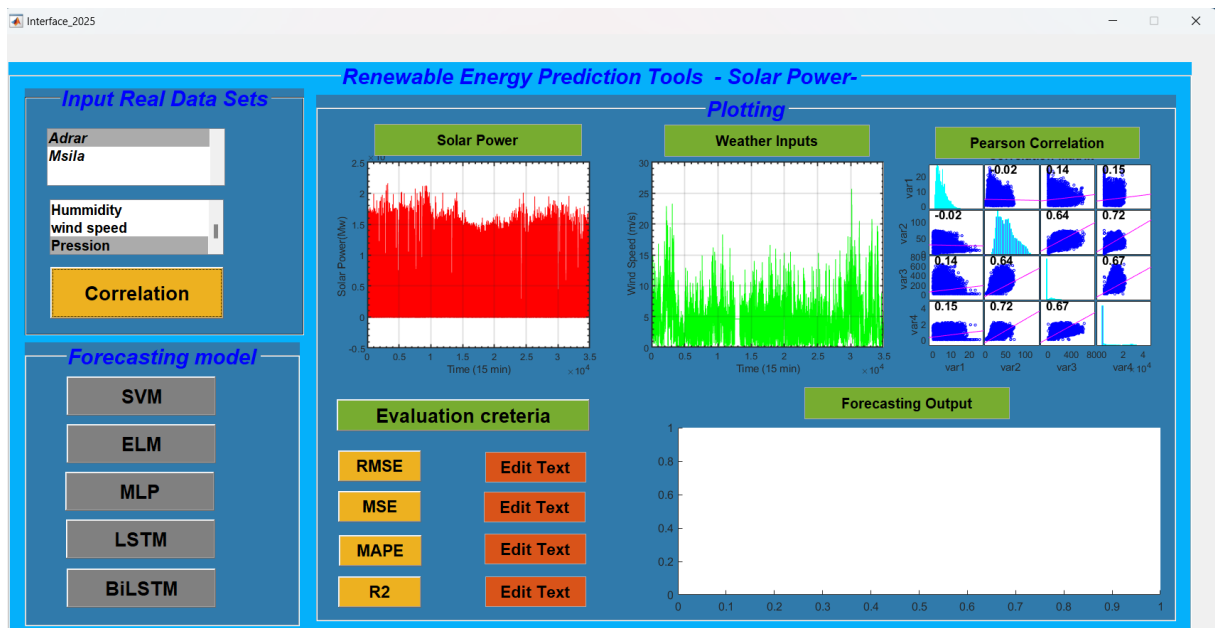


Figure 3.17 : Visualisation matrice de corrélation de Pearson entre les différentes variables météorologiques et la puissance solaire.

### 3.7.2.3 Bloc de Choix du Modèle de Prédiction

Ce bloc permet à l'utilisateur de choisir l'algorithme d'intelligence artificielle à utiliser pour la prédiction de la puissance solaire, en fonction des performances souhaitées ou de la nature des données. Cette fonctionnalité est représentée par la partie 7 sur la figure 3.15.

Plusieurs modèles sont proposés dans l'interface, parmi lesquels : SVM (Support Vector Machine), MLP (Multi-Layer Perceptron), ELM (Extreme Learning Machine), LSTM (Long Short-Term Memory) et BiLSTM (Bidirectional LSTM)

Chaque bouton de prédicteur déclenche l'exécution du modèle correspondant et produit une prédiction basée sur les variables météorologiques sélectionnées. Ce bloc constitue une étape centrale dans le processus de modélisation.

### 3.7.2.4 Bloc Résultats de la Prédiction

Ce bloc affiche les résultats numériques et graphiques de la prédiction réalisée. Il contient les parties 8 et 9 illustrées par la figure 3.15 :

- **Partie 8 : Évaluation des performances** correspondant à l'élément 8 sur la figure globale de l'interface présente les indicateurs statistiques utilisés pour quantifier la précision du modèle de prédiction, tels que le RMSE, MSE, MAPE et le coefficient de détermination  $R^2$ . Ces valeurs numériques permettent de comparer objectivement les performances des différents modèles sélectionnés.

- **Partie 9 : Courbe de sortie de prédiction** correspondant à l'élément 9 sur la figure globale de l'interface affiche un graphique représentant la comparaison entre les valeurs réelles de la puissance solaire et celles prédites par le modèle sélectionné. Cette visualisation permet d'évaluer visuellement la performance du modèle et de détecter les éventuels écarts de prédiction.

La figure 3.18 illustre la courbe de la puissance solaire réelle comparée à celle prédite par le modèle sélectionné (SVM), accompagnée des valeurs numériques des métriques de performance (RMSE, MSE, MAPE,  $R^2$ ), permettant ainsi d'évaluer la qualité de la prédiction.

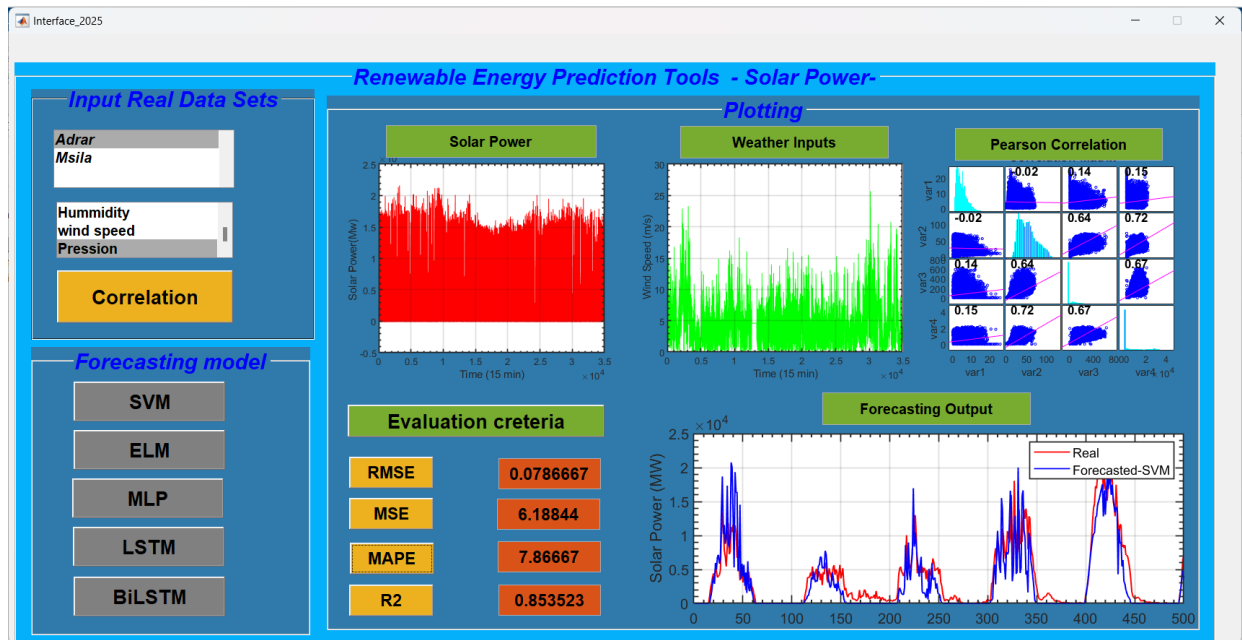


Figure 3.18 Visualisation de la courbe réelle et prédite de la puissance solaire et affichage des valeurs numériques des métriques de performance.

### 3.7.4 Fonctionnalités principales

Cette interface propose plusieurs fonctionnalités essentielles qui la rendent opérationnelle et efficace pour les tâches de prévision :

- **Chargement des données** : l'utilisateur peut importer un fichier de données météorologiques (Excel ou CSV) correspondant à un site donné. Ces données sont automatiquement prétraitées et préparées pour l'analyse.
- **Sélection et affichage des variables** : grâce à une liste déroulante, l'utilisateur peut sélectionner les variables d'entrée. Les courbes associées sont affichées dynamiquement pour permettre une première analyse visuelle.
- **Analyse de la corrélation** : un bouton « Corrélation » déclenche le calcul des coefficients de corrélation entre les variables météorologiques et la puissance solaire, affichés sous forme de matrice graphique.

- **Choix du modèle prédictif** : un simple clic sur l'un des boutons (SVM, ELM, MLP, LSTM, BiLSTM) permet d'entraîner le modèle correspondant avec les données sélectionnées, puis d'obtenir une prédiction.
- **Affichage des résultats** : le graphique de sortie affiche la comparaison entre les valeurs réelles et les valeurs prédites. Les métriques de performance sont également mises à jour automatiquement.

### 3.8. Conclusion

Dans ce chapitre, nous avons présenté la conception et le développement d'une interface utilisateur sous MATLAB, dédiée à l'installation optimale de sites d'énergies renouvelables. Cette interface permet d'intégrer de manière efficace les données techniques, économiques et environnementales nécessaires à la prise de décision. L'étude de cas abordée a permis d'illustrer la pertinence de l'outil, en simulant différentes configurations d'implantation et en évaluant leur faisabilité.

Une attention particulière a été portée à la structuration de la base de données, qui regroupe les informations géographiques, météorologiques, économiques ainsi que les caractéristiques techniques des différentes technologies (solaire, éolien, etc.). Cette base constitue le cœur du système, assurant la cohérence et la fiabilité des résultats obtenus.

Par ailleurs, des critères d'évaluation multicritères ont été intégrés dans l'interface afin d'aider à la sélection des sites optimaux. Ces critères incluent notamment le potentiel énergétique, le coût d'installation, l'impact environnemental, ainsi que la proximité aux infrastructures existantes.

En somme, cette interface MATLAB représente un outil d'aide à la décision flexible et évolutif, capable de s'adapter à divers contextes territoriaux et technologiques. Elle constitue une contribution concrète à la planification rationnelle des projets d'énergie renouvelable.

### *Conclusion générale*

Face aux défis énergétiques actuels et à la nécessité de transition vers des sources d'énergie durables, l'optimisation de l'implantation des sites de production d'énergies renouvelables constitue une étape stratégique. Dans ce contexte, notre projet s'est inscrit dans une démarche innovante visant à concevoir une interface MATLAB capable d'assister les décideurs, ingénieurs et planificateurs dans le choix optimal des emplacements pour les installations solaires et éoliennes.

L'interface développée combine des fonctionnalités avancées de traitement de données, de visualisation graphique et de simulation, permettant une analyse approfondie de plusieurs scénarios d'implantation. Elle intègre une base de données riche, structurée et évolutive, regroupant des informations climatiques, économiques, topographiques et techniques. Cette base constitue un socle essentiel pour garantir la fiabilité et la pertinence des résultats générés.

L'étude de cas réalisée a démontré l'utilité pratique de l'interface, mettant en évidence sa capacité à fournir une aide à la décision rigoureuse, rapide et accessible. En facilitant l'analyse et la synthèse d'une grande quantité d'informations complexes, cette interface MATLAB offre une solution performante et adaptable aux besoins des acteurs du secteur des énergies renouvelables.

En conclusion, ce projet apporte une contribution concrète à la digitalisation de la planification énergétique, tout en ouvrant la voie à de futures améliorations telles que l'intégration d'algorithmes d'intelligence artificielle, l'extension à d'autres types de ressources (biomasse, hydraulique), ou encore la connexion à des systèmes d'information géographique (SIG) pour un rendu plus précis et interactif.

## Références :

### Chapitre 01 :

- [1] Nick Jelley « Les énergies renouvelables. ». P 9.10
- [2] Connaissance des Énergies : le média pédagogique sur l'énergie. Consulté le 2 mars 2025, sur <https://www.connaissancedesenergies.org/>
- [3] UNFCCC. Les énergies renouvelables. Consulté le 12 avril 2025, sur : <https://unfccc.int/fr/news/les-energies-renouvelables>
- [4] [actu-environnement.com/ae/dictionnaire\\_environnement/definition/energie\\_hydraulique](https://actu-environnement.com/ae/dictionnaire_environnement/definition/energie_hydraulique).
- [5] Organisation des Nations Unies. Les objectifs de développement durable. Consulté le 19 avril 2025, sur <https://www.nationsunies.fr/odd/>
- [6] Thème : énergiehydraulique. Présenter par : Boudgomia Hicham, Cherefrabah Baranti mahdi. Encadré par : Mr. Dahak .M
- [7] Dictionnaire de français Larousse [archive] », sur [larousse.fr](http://larousse.fr) (consulté le 8 mai 2022 ).
- [8] Le terme « biomasse » est apparu en 1966. Il est composé de masse avec le préfixe bio-, du grec ancien bíos, « vie ». [www.cnrtl.fr](http://www.cnrtl.fr) (consulté le 8 mai 2022).
- [9] Williams, Dahiya et Porter 2020, P. 7.
- [10] National Geographic – Toute l'actualité scientifique au quotidien. Consulté le 12 avril 2025, sur : <https://www.nationalgeographic.fr/>
- [11] ADEME (Agence de la Transition Écologique). (2021).
- [12] Operaenergie par caroline dusanter.
- [13] Thèse de doctorat de l'université de Toulouse, Janvier 2009.
- [14] Connaissance des énergies.
- [15] PROJET FIN D'ETUDE.ENERGIE SOLAIRE PHOTOVOLTAIQUE ET SYSTEME PHOTOVOLTAIQUE AUTONOME.PAR : TOUMI Nihed et KOUDA Iskandar. Monsieur ADJABI Mohamed.
- [16] [edf-solutions-solaires](http://edf-solutions-solaires).
- [17] DualSun – Solutions solaires photovoltaïques et thermiques. Consulté le 18 janvier 2025, sur : <https://dualsun.com/>
- [18] Wikipédia, l'encyclopédie libre. Consulté le 14 mars 2025, sur : <https://fr.wikipedia.org/>
- [19] Wikipédia. Énergie solaire thermique. Consulté le 17 Février 2025, sur : [https://fr.wikipedia.org/wiki/%C3%89nergie\\_solaire\\_thermique](https://fr.wikipedia.org/wiki/%C3%89nergie_solaire_thermique).
- [20] [Étude comparative du Ministère fédéral allemand de l'environnement (BMU) [archive], citée par Terre et Finance].
- [21] [Claire Vaille, « Les avantages du solaire thermique, histoire et perspectives », numéro 447, sur Bulletin ADIT BE Allemagne, Ambassade de France en Allemagne/ADIT, 30 juillet 2009].

- [22] ADEME (Agence de la transition écologique), Guide pratique sur le solaire combiné (2021).
- [23] Installation de panneaux solaires | Otovo. Consulté le 19 février 2025, sur :  
<https://www.otovo.fr/>
- [24] L'Énergie Tout Compris. Solaire thermodynamique : fonctionnement et types de centrales. Consulté le 20 février 2025, sur :  
<https://www.lenergiesoutcompris.fr/actualites-conseils/solaire-thermodynamique-fonctionnement-et-types-de-centrales-48118>
- [25] Wikipédia. Énergie solaire en France. Consulté le 20 Février mars , sur :  
[https://fr.wikipedia.org/wiki/Énergie\\_solaire\\_en\\_France](https://fr.wikipedia.org/wiki/Énergie_solaire_en_France)
- [26] Eco Delta. Quel futur pour l'énergie solaire. Consulté le 15 mars 2025, sur :  
<https://ecodelta.fr/quel-futur-pour-lenergie-solaire>

## Chapitre 02 :

- [27] Formations Data & Bootcamps : formez-vous à la Data avec DataBird !. Consulté le 18 mai 2025, sur :  
<https://www.data-bird.co/>
- [28] *"Pattern Recognition and Machine Learning"* —Christopher Bishop.
- [29] Mémoire de Fin d'études Master. Présenté par : Mohamed AbdElmoumen DJABALLAH. Encadré Par : Dr. Ouarda ZEDADRA.
- [30] Régis Bourbonnais, Michel Terraza, Analyse des séries temporelles : applications à l'économie et à la gestion, 4e Edt. Dunod, 2016.
- [31] Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
- [32] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). *Gradient-based learning applied to document recognition*. Proceedings of the IEEE, 86(11), 2278–2324.
- [33] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- [34] Cormack, G. V. (2007). *Email spam filtering: A systematic review*. Foundations and Trends in Information Retrieval, 1(4), 335–455.
- [35] Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Pre
- [36] Jain, A. K. (2010). *Data clustering: 50 years beyond K-means*. Pattern Recognition Letters, 31(8), 651–666.
- [37] Hinton, G. E., & Salakhutdinov, R. R. (2006). *Reducing the dimensionality of data with neural networks*. Science, 313(5786), 504–507
- [38] Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). *Isolation forest*. In *2008 Eighth IEEE International Conference on Data Mining* (pp. 413–422). IEEE

- [39] Zhu, X. (2005). *Semi-supervised learning literature survey*. University of Wisconsin-Madison, Computer Sciences TR 1530
- [40] Nigam, K., McCallum, A., Thrun, S., & Mitchell, T. (2000). *Text classification from labeled and unlabeled documents using EM*. *Machine Learning*, 39(2), 103–134.
- [41] Chapelle, O., Schölkopf, B., & Zien, A. (2006). *Semi-Supervised Learning*. MIT Press.
- [42] Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. *Synthesis Lectures on Human Language Technologies*, 5(1), 1–167.
- [43] LeCun, Y. (2021). *Self-supervised learning: The dark matter of intelligence*. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*.
- [44] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. In *Proceedings of NAACL-HLT*, 4171–4186.
- [45] Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). *A Simple Framework for Contrastive Learning of Visual Representations*. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*.
- [46] Baevski, A., Zhou, H., Mohamed, A., & Auli, M. (2020). *wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations*. In *Advances in Neural Information Processing Systems (NeurIPS)*. P 9.
- [47] Russell S., Norvig P. 2002. Intelligence artificielle : une approche moderne.
- [48] Pearce CB, Gunn SR, Ahmed A., Johnson CD. L'apprentissage automatique peut améliorer la prédiction de la gravité de la pancréatite aiguë en utilisant les valeurs d'admission du score APACHE II et de la protéine C-réactive. *Pancréatologie*. 2006 ; 6 : 123–131.doi.
- [49] Sabanayagam C., Xu D., Ting DSW, Nusinovici S., Banu R., Hamzah H., et al. Un algorithme d'apprentissage profond pour détecter l'insuffisance rénale chronique à partir de photographies rétinienne dans les populations locales. *Lancet Digit Health*. 2020.
- [50] Fu Y., Jung AW, Torne RV, Gonzalez S., Vöhringer H., Shmatko A., et al. L'histopathologie computationnelle pancréatique révèle les mutations, la composition tumorale et le pronostic. *Nat Cancer*. 2020.
- [51] Rezazade Mehrizi MH, van Ooijen P., Homan M. Applications de l'intelligence artificielle (IA) en radiologie diagnostique : une étude technographique. *EurRadiol*. 2021.
- Wikipédia. Perceptron multicouche. Consulté le 18 mars 2025, sur : [https://fr.wikipedia.org/wiki/Perceptron\\_multicouche](https://fr.wikipedia.org/wiki/Perceptron_multicouche)
- [53] Frank Rosenblatt, « The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain », *Psychological Review*, vol. 65, no 6, 1958, p. 386–408.
- [54] IBM France – Leader mondial du Cloud hybride et de l'IA. Consulté le 5 avril 2025, sur : <https://www.ibm.com/fr-fr/>
- [55] Wikipédia, l'encyclopédie libre. Consulté le 28 mars 2025, sur : <https://fr.wikipedia.org/>

- [56] Scribd – la bibliothèque numérique mondiale. Consulté le 11 janvier 2025, sur : <https://fr.scribd.com/>
- [57] La définition « sans programmation explicite » est souvent attribuée à Arthur Samuel , inventeur du terme « apprentissage automatique » en 1959.
- [58] IBM. 22 septembre 2021. Archivé de l'original le 27 décembre 2023. Consulté le 27 juin 2023.
- [59] Hu, Junyan ; Niu, Hanlin ; Carrasco, Joaquin ; Lennox, Barry ; Arvin, Farshad (2020). « Exploration autonome multi-robots basée sur Voronoi dans des environnements inconnus via l'apprentissage par renforcement profond » (PDF). Transactions IEEE sur la technologie véhiculaire.
- [60] L'apprentissage automatique et la reconnaissance de formes « peuvent être considérés comme deux facettes d'un même domaine ». Friedman, Jerome H.(1998). « Exploration de données et statistiques : quel est le lien ? ». Informatique et statistiques.
- [61] «<https://www.journaldunet.fr/web-tech/guide-de-l-intelligence-artificielle/1501879-machine-a-vecteurs-de-support-svm-definition-et-cas-d-usage> (Consulted on 10/ 05/ 2025) (cf. p. 30).
- [62] IBM+1Wikipédia+1.
- [63] IBM.
- [64] Medium+3IBM+3Analytics Vidhya+3.
- [65] «<https://medium.com/analytics-vidhya/machine-learning-support-vector-regression181aea35bedf>,(Consulted on 03/ 05/ 2025), » (cf. p. 32).
- [66] "Deep learning" : les dessous d'une technologie de rupture [archive], analyse prospective, Futurible.
- [67] ARNOLD, Ludovic, REBECCHI, Sébastien, CHEVALLIER, Sylvain, et al. An introduction to deep learning. In: European Symposium on Artificial Neural Networks (ESANN). 2011.
- [68] "Understanding LSTM Networks -- colah's blog." [Online]. Available: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>. [Accessed: 08-Jun-2018].
- [69] DataScientest – école en ligne en Data, Tech & Cybersécurité. Consulté le 13 Février 2025, sur : <https://datascientest.com/>
- [70] ScienceDirect.
- [71] GeeksforGeeks.
- [72] NACER, Saadi Mohamed, NADIA, Bouteraa, ABDELGHANI, Redjati, ET MOHAMED, Boughazi. A novel method for bearing fault diagnosis based on BiLSTM neural networks. The International Journal of Advanced Manufacturing Technology, 2023, p. 1-16.
- [73] Mémoire Présenté en vue de l'obtention du Diplôme de Master. Thème : Diagnostic des défauts de roulement basé sur les réseaux de neurones BiLSTM. Présenté par : - REDJATI Meriem Takoua - RACHID Kaouter. Encadrant : BOUGHAZI Mohamed.
- [74] GeeksforGeeks – A Computer Science Portal for Geeks. Consulté le 22 mars 2025, sur :

<https://www.geeksforgeeks.org/>

### **Chapitre 03 :**

[75] Mémoire master. Présenté par :REDJATI Meriem Takoua RACHID Kaouter. Encadrant: BOUGHAZI Mohamed.

[76] Cecil Kuhne, « Near Death in the Desert [archive] », sur Google Books, KnopfDoubledayPublishing Group, 8 juin 2011.

[77] Wikipédia, l'encyclopédie libre. Consulté le 30 avril 2025, sur :

<https://fr.wikipedia.org/>

[78] Wikipédia, l'encyclopédie libre. Consulté le 24 mai 2025, sur :

<https://fr.wikipedia.org/>

[79] Climate-Data.org – Données climatiques mondiales. Consulté le 02 avril 2025, sur :

<https://fr.climate-data.org/>