



République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Ahmed Draïa Adrar
Faculté des Sciences des Matériaux, Mathématiques et Informatique



Thèse

Présentée en vue de l'obtention du diplôme de Doctorat 3^{ème} cycle en
Sciences des Matériaux
Option : Physique des rayonnements

Par :

SENOUCI Ahlam

Etude et modélisation par méta-heuristique hybride méthode pour l'estimation du rayonnement en climat saharien

Soutenue publiquement le : 01 / 07 / 2025 Devant le jury :

Président	ARBAOUI Iliace	MCA	Univ Adrar
Directeur de thèse	BENATIALLAH Ali	Pr.	Univ Adrar
Co-Directeur de thèse	BOUCHOUICHA Kada	DR.	CDER Alger
Examineur	BENATIALLAH Djelloul	Pr.	Univ Adrar
Examineur	BOUDAUD Lahouaria	Pr.	Univ Adrar
Examineur	MAOUEDJ Rachid	DR.	URER MS. Adrar

Année Universitaire : 2024/2025

Remerciements

Je tiens à exprimer ma profonde gratitude à toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce travail de recherche.

À mon directeur de thèse, le Professeur Ali BENATIALLAH,

Merci pour votre encadrement précieux, vos conseils avisés et votre disponibilité tout au long de ce parcours. Votre expertise et votre rigueur scientifique ont été une source d'inspiration et m'ont guidé à chaque étape de cette recherche.

Au Professeur Maouedj Rachid,

Je vous suis reconnaissant pour vos remarques constructives et votre soutien lors des différentes phases de ce travail. Vos suggestions ont grandement contribué à l'amélioration de ce travail.

Je remercie sincèrement **Pr. Djelloul Benatallah**, qui m'a aidé et m'a donné le fil conducteur nécessaire pour que je puisse entamer et mener à bien ce travail.

Aux membres respectés du jury

Docteur Iliace ARBAOUI, les Professeurs Djelloul BENATIALLAH, Houaria BOUDAUD, Kada BOUCHOUICHA,

Je vous remercie chaleureusement pour le temps que vous avez consacré à l'évaluation de ce travail, ainsi que pour vos critiques et recommandations qui ont permis d'enrichir cette étude.

Enfin, je n'oublie pas tous ceux qui, de près ou de loin, m'ont encouragé et soutenu dans cette aventure académique.

Avec toute ma reconnaissance,

Ahlam Senouci

Dédicace

À mes parents bien-aimés,

Vous qui avez été mon rocher, ma lumière et mon soutien inconditionnel à chaque étape de ce parcours. Votre amour, vos sacrifices et vos encouragements constants ont fait de moi la personne que je suis aujourd'hui. Cette réussite est autant la vôtre que la mienne.

À mon cher époux, Mohammed Mounir,

Mon partenaire de vie, ma force tranquille. Merci pour ta patience infinie, ton soutien sans faille et ton amour qui m'ont portée dans les moments les plus exigeants de cette aventure.

À mes anges, Maram et Majed,

Mes rayons de soleil, ma plus grande fierté. Vous m'avez donné la force d'avancer même lorsque le chemin semblait difficile. Que cette réalisation vous inspire à toujours croire en vos rêves.

À mes sœurs et frères, Haifaa, Baraa, Afnane, Younes, Fakher Eddine et Ayoub, Merci pour votre affection, votre complicité et vos encouragements précieux. Vous avez été ma deuxième colonne vertébrale, toujours présents pour me soutenir et me motiver.

Cette thèse est le fruit de notre histoire familiale, tissée d'amour, de persévérance et de foi mutuelle. Chacune de ces pages porte un peu de vous tous.

Avec tout mon amour et ma reconnaissance éternelle,

Résumé

Dans cette thèse, nous avons développé et évalué plusieurs approches intelligentes pour l'estimation du rayonnement solaire quotidien dans un climat saharien, caractérisé par des conditions extrêmes et une forte variabilité. L'objectif principal est d'améliorer la précision des prévisions afin de favoriser l'intégration des énergies renouvelables, en particulier l'énergie solaire, dans ces régions.

Nous avons d'abord appliqué des modèles classiques d'apprentissage comme les réseaux de neurones artificiels (ANN) et les systèmes neuro-flous adaptatifs (ANFIS), couplés à deux algorithmes d'optimisation méta-heuristiques : l'algorithme génétique (GA) et l'optimisation par essaim particulaire (PSO). Ces combinaisons (ANN-GA, ANN-PSO, ANFIS-GA, ANFIS-PSO) ont permis une amélioration notable de la performance des modèles, notamment en termes de précision et de robustesse.

Ensuite, une nouvelle méthode hybride avancée a été proposée : EEMD-GA-LSTM, combinant le débruitage des séries temporelles par EEMD (Ensemble Empirical Mode Decomposition), l'optimisation par GA, et la puissance prédictive des réseaux LSTM (Long Short-Term Memory). Cette approche a démontré des résultats très prometteurs, surpassant les modèles classiques, en particulier dans la capture des tendances complexes et non linéaires du rayonnement solaire.

Les simulations ont été réalisées sur des données réelles collectées dans quatre régions sahariennes algériennes : Biskra, Adrar, Tamanrasset et Tindouf. Les performances des modèles ont été évaluées à l'aide d'indicateurs statistiques standards (RMSE, MAE, R). Les résultats obtenus montrent que la méthode EEMD-GA-LSTM offre une estimation plus fiable et adaptée aux conditions climatiques extrêmes du Sahara.

Mot clé : Rayonnement solaire, Climat saharien, Méta-heuristique, Réseaux de neurones (ANN), Systèmes neuro-flous (ANFIS), Algorithme génétique (GA), Optimisation par essaim particulaire (PSO), EEMD, LSTM, EEMD-GA-LSTM.

Abstract

In this thesis, we have developed and evaluated several intelligent approaches for the estimation of daily solar radiation in a Saharan climate, characterized by extreme conditions and high variability. The main objective is to improve forecasting accuracy to support the integration of renewable energies, particularly solar energy, in these regions.

We first applied conventional machine learning models such as Artificial Neural Networks (ANN) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS), combined with two metaheuristic optimization algorithms: Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). These combinations (ANN-GA, ANN-PSO, ANFIS-GA, ANFIS-PSO) significantly enhanced model performance, especially in terms of accuracy and robustness.

Subsequently, an advanced hybrid method was proposed: EEMD-GA-LSTM, which integrates time series denoising through Ensemble Empirical Mode Decomposition (EEMD), optimization via GA, and the predictive power of Long Short-Term Memory (LSTM) neural networks. This approach showed very promising results, outperforming conventional models, particularly in capturing complex and nonlinear trends in solar radiation.

Simulations were conducted using real-world data collected from four Saharan regions of Algeria: Biskra, Adrar, Tamanrasset, and Tindouf. The models' performance was evaluated using standard statistical indicators (RMSE, MAE, R). The results demonstrate that the EEMD-GA-LSTM method provides a more reliable estimation, well-suited to the extreme climatic conditions of the Sahara.

Keywords: Solar radiation, Saharan climate, Metaheuristic, Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Genetic Algorithm (GA), Particle Swarm Optimization (PSO), EEMD, LSTM, EEMD-GA-LSTM.

الملخص

في هذه الأطروحة، قمنا بتطوير وتقييم عدة مقاربات ذكية لتقدير الإشعاع الشمسي اليومي في مناخ صحراوي يتميز بظروف مناخية قاسية وتفاوت كبير. الهدف الرئيسي هو تحسين دقة التنبؤات من أجل دعم إدماج الطاقات المتجددة، وخاصة الطاقة الشمسية، في هذه المناطق.

في البداية، قمنا بتطبيق نماذج تعلم تقليدية مثل الشبكات العصبية الاصطناعية والأنظمة العصبية الضبابية التكيفية، مقترنة بخوارزميات تحسين ذكية مثل الخوارزمية الجينية وخوارزمية تحسين سرب الجسيمات. وقد سمحت هذه التركيبات بتحسين ملحوظ في أداء النماذج، لا سيما من حيث الدقة والموثوقية.

بعد ذلك، تم اقتراح طريقة هجينة متقدمة تجمع بين إزالة التشويش من البيانات باستخدام تقنية التحليل التجريبي النمطي المجمع، والتحسين بواسطة الخوارزمية الجينية، والقدرة التنبؤية لشبكات الذاكرة طويلة وقصيرة الأمد. وقد أظهرت هذه الطريقة نتائج واعدة جداً، حيث تفوقت على النماذج التقليدية، خصوصاً في تمثيل الأنماط المعقدة وغير الخطية للإشعاع الشمسي.

وقد تم إجراء المحاكاة بالاعتماد على بيانات حقيقية مأخوذة من أربع مناطق صحراوية جزائرية: بسكرة، أدرار، تمنراست، وتندوف. وتم تقييم أداء النماذج باستخدام مؤشرات إحصائية معيارية مثل جذر متوسط مربع الخطأ، ومتوسط الخطأ المطلق، ومعامل التحديد. وتُظهر النتائج أن الطريقة المقترحة توفر تقديرًا أكثر موثوقية وتكيفًا مع الخصائص المناخية القاسية للصحراء.

الكلمات المفتاحية : الإشعاع الشمسي، المناخ الصحراوي، خوارزميات ذكية، الشبكات العصبية، الأنظمة الضبابية التكيفية، الخوارزمية الجينية، خوارزمية تحسين السرب، التحليل التجريبي النمطي المجمع، شبكات الذاكرة.

Tables des matières

Résumé	iii
Liste des figures	vi
Liste des tableau	viii
Nomenclature	ix
Introduction générale	02

Chapitre I : Énergie Solaire : Origine, Mesure et Influences

Introduction	05
1. Le soleil	05
2. Le système solaire	07
3. Le bilan radiatif terrestre	08
3.1. Paramètres géographiques	08
3.1.1. La latitude	08
3.1.2. La longitude	08
3.1.3. L'altitude	08
3.2. Paramètres astronomiques	09
3.2.1. Réparage du soleil	09
3.2.2. Déclinaison de soleil	10
3.2.3. Angle horaire du soleil	10
3.3. Cordonnées horizontales	11
3.3.1. La hauteur du soleil	11
3.3.2. Azimute du soleil	12
4. Les temps solaire	12
4.1. Le temps solaire moyen	12
4.2. Le temps universel	13
4.3. Le temps légal	13
4.4. Le temps solaire vrai	13
4.5. L'équation de temps	13
4.6. Transition du temps légal vers le temps solaire vrai	14
5. Émission du soleil	15
5.1. Nature de rayonnement solaire	15

Tables des matières

5.2.	Propagation des ondes électromagnétiques	15
5.3.	Propriété des ondes électromagnétiques	15
5.4.	Spectre du soleil	16
6.	Le rayonnement solaire	16
6.1.	Le couple terre soleil	16
6.1.1.	Orbite et saison	16
6.1.2.	Lumière et Energie	17
6.2.	Climat et météorologie	17
6.3.	Espace et exploration	17
6.4.	Correction de la distance terre soleil	17
6.5.	La constante solaire	18
6.6.	Correction de la constante solaire	18
6.7.	Flux solaire d'un plan horizontale	19
6.8.	Irradiation journalière sur un plan horizontale	19
7.	Mesure et analyse du rayonnement solaire	20
7.1.	Technique de mesure de rayonnement solaire	20
7.1.1.	Pyranomètres	20
7.1.2.	Pyrhéliomètres	20
7.1.3.	Albédo-mètre	21
7.1.4.	Héliographes	21
7.2.	Utilisation des données satellitaire	22
7.2.1.	METEOSAT	22
7.2.2.	EUMETSAT	22
7.2.3.	Hélio-Clim	22
8.	Impacte des conditions Atmosphériques	22
8.1.	Nébulosité	22
8.2.	Pollution atmosphériques	23
8.3.	Vapeur d'eau	23
9.	Bases de données climatiques	23
10.	Modélisation et simulation	23

Tables des matières

Conclusion	24
Références	24

Chapitre II : Estimation du Rayonnement Solaire dans les Régions Arides

Introduction	28
1. Rayonnement solaire une source énergétique majeur	28
1.1. Définition et composantes de rayonnement solaire	29
1.2. Le rayonnement global	29
2. Facteurs influencent le rayonnement solaire	30
3. Le rayonnement solaire: un moteur de la vie sur terre	32
4. Les méthodes traditionnel d'estimation du rayonnement solaire	32
4.1. modèles empiriques	33
4.2. modèles physiques	35
4.3. modèles basé sur IA	36
5. Spécificités du climat saharien et impact sur l'estimation du rayonnement solaire	40
5.1. Caractéristique climatique de Sahara	40
5.2. Conséquences sur le rayonnement solaire	40
6. Difficulté liées a l'estimation de rayonnement solaire dans le Sahara	41
7. Les limites des modèles existants dans le contexte saharien	42
8. Les méta-heuristiques : une approche prometteuse	44
9. Exploration efficace de l'espace de recherche	45
Conclusion	47
Références	47

Chapitre III : Méthodologie et outils

Introduction	51
1. Collecte et prétraitement des donné météorologique	51
1.1. Données météorologiques	51
1.2. Données satellitaires	52
1.3. Autres données	52

Tables des matières

1.4. Prétraitement des données	52
2. Description des données	53
3. Méthodes de prévision	53
3.1. ANN	53
3.2. ANFIS	54
3.3. LSTM	54
3.4. Algorithme génétique	55
3.5. Optimisation par essaim particule	55
3.6. Décomposition par EEMD et couplage avec LSTM-GA	56
4. Évaluation des performances des modèles	56
4.1. R	56
4.2. RMSE	57
4.3. MAE	57
5. Outils et environnement de travail	57
5.1. Matlab	57
5.2. Python	58
Conclusion	59
Références	59

Chapitre IV: Modélisation et Prédiction du Rayonnement Solaire Global par des Méthodes métaheuristique

Introduction	62
1. Sélection des sites et collecte des données	63
1.1. Caractéristiques géographiques des sites	64
1.2. Données météorologiques collectées	64
1.3. Périodes de collecte des données	65
1.4. Validation des données	65
2. Prétraitements des données	66
2.1. Validation des données	66
2.2. Normalisation des données	66
2.3. Techniques de rééchantillonnage	67

3. Développent des modèles hybrides	68
3.1. ANN	68
3.2. ANFIS	69
3.3. Optimisation par algorithme métaheuristique	69
4. Évaluation des performances	70
5. Algorithme	71
6. Résultat et discussion	73
Conclusion	77
Références	78

Chapitre V : Estimation Innovante du Rayonnement Solaire Quotidien : Méthode EEMD-GA-LSTM

Introduction	82
1. Présentation détaillée de la méthode hybride (EEMD-GA-LSTM)	83
1.1. Introduction a l'approche hybride	83
1.2. Décomposition des séries temporelle avec EEMD	83
1.3. Optimisation des paramètres avec GA	84
1.4. Modèle de prédiction avec LSTM	84
1.5. Avantage de la méthodes EEMD-GA-LSTM	85
2. Justification de choix de l'approche EEMD-GA-LSTM	86
2.1. Complexité et non linéaire des données météorologiques	86
2.2. Optimisation des hyperparamètres pour une meilleur précision	87
2.3. Capture les dépendances temporelle complexe	87
2.4. Traitement efficace des données limitées	87
2.5. Robustesse et généralisation	88
2.6. Comparaison avec d'autre méthodes	88
3. Méthodologie de mise en œuvre de la méthodes EEMD-GA-LSTM	89
4. Validation croisées et évaluation des performances	92
5. Algorithme de prédiction du rayonnement solaire basées sur EEMD-GA-LSTM	96
6. Résultat et discussions	99

Tables des matières

Conclusion	110
Références	111
Conclusion générale	116

Liste des figures

Chapitre I : Énergie Solaire : Origine, Mesure et Influences

Figure 1 : la distribution spectrale du rayonnement solaire reçu par la terre	06
Figure 2 : Mouvement du globe terre par rapport au soleil.	07
Figure 3 : Coordonnées géographiques du lieu.	09
Figure 4 : Repère équatorial.	09
Figure 5 : Variation de l'angle horaire en fonction du temps.	10
Figure 6: cordonnées horizontales.	11
Figure 7 : Variation de la hauteur du soleil au cours d'une année.	12
Figure 8 : Equation du temps et déclinaison en fonction du numéro de jour de l'année.	14
Figure 9 : Spectre de rayonnement solaire	16
Figure 10 : Évolution annuelle de la constante solaire.	19
Figure 11 : Pyranomètres	21
Figure 12 : Pyrhéliomètres	21
Figure 13 : Albédo-mètre	22
Figure 14 : Héliographes	22

Chapitre II : Estimation du Rayonnement Solaire dans les Régions Arides

Figure 1: Les différents composants du rayonnement solaire.	30
Figure 2: l'altitude.	31
Figure 3 : Structure de base d'un MLP	37
Figure 4: Illustration des couches ANFIS avec entrées floues et sortie précise.	39

Chapitre IV: Modélisation et Prédiction du Rayonnement Solaire Global par des Méthodes métaheuristique

Figure 1 : Irradiation globale reçue sur une surface horizontale	63
Figure 2: Organigramme du Modèle ANN basé sur GA et PSO	71
Figure 3: Organigramme du Modèle ANFIS optimisé par GA et PSO	72
Figure 4: Comparaison entre l'irradiation calculée par les modèles ANN, ANN-GA, ANN-PSO, ANFIS, ANFIS-GA et ANFIS-PSO et les valeurs mesurées à Adrar.	73
Figure 5: Comparaison entre l'irradiation calculée par les modèles ANN, ANN-GA, ANN-PSO, ANFIS, ANFIS-GA et ANFIS-PSO et les valeurs mesurées à Tindouf.	74
Figure 6: Comparaison entre l'irradiation calculée par les modèles ANN, ANN-GA, ANN-PSO, ANFIS, ANFIS-GA et ANFIS-PSO et les valeurs mesurées à Tamanrasset.	75

Chapitre V : Estimation Innovante du Rayonnement Solaire Quotidien : Méthode EEMD-GA-LSTM

Figure 1: Schéma de modèle proposé	98
------------------------------------	----

Liste des figures

Figure 2 : Signal de Rayonnement Solaire Normalisé dans la région de Biskra	99
Figure 3 : Résultats de la décomposition EEMD dans la région de Biskra.	100
Figure 4 : Signal de Rayonnement Solaire Normalisé dans la région de Adrar.	100
Figure 5 : Résultats de la décomposition EEMD dans la région d'Adrar.	101
Figure 6 : Signal de Rayonnement Solaire Normalisé dans la région de Tamanrasset.	101
Figure 7 : Résultats de la décomposition EEMD dans la région de Tamanrasset.	102
Figure 8 : Signal de Rayonnement Solaire Normalisé dans la région de Tindouf.	102
Figure 9: Résultats de la décomposition EEMD dans la région de Tindouf.	103
Figure 10 : Changement global du RMSE dans les régions de Biskra, Adrar, Tamanrasset et Tindouf.	104
Figure 11 : Changement global du MAE dans les régions de Biskra, Adrar, Tamanrasset et Tindouf.	105
Figure 12 : Changement global du coefficient de détermination (R^2) dans les régions de Biskra, Adrar, Tamanrasset et Tindouf.	106
Figure 13 : Irradiation globale observée vs prédite en utilisant l'approche EEMD-GA-LSTM dans les régions de Biskra, Adrar, Tamanrasset et Tindouf	107
Figure 14 : Irradiation prédite en utilisant l'approche EEMD-GA-LSTM en fonction de l'irradiation actuelle dans les régions de Biskra, Adrar, Tamanrasset et Tindouf.	108

Chapitre II : Estimation du Rayonnement Solaire dans les Régions Arides

Tableau 1: Comparaison des Modèles IA vs. Empiriques/Physiques	40
--	----

Chapitre IV: Modélisation et Prédiction du Rayonnement Solaire Global par des Méthodes métaheuristique

Tableau1 : Coordonnées géographiques des trois sites étudiés situés en zone aride	64
Tableau 2: Scores statistiques	76

Chapitre V : Estimation Innovante du Rayonnement Solaire Quotidien : Méthode EEMD-GA-LSTM

Tableau1 : Résultat des indicateurs statistiques.	109
Tableau 2 : Comparaison des performances des modèles présentés avec les résultats de diverses études.	109

Nomenclature

Liste des symboles

Angles et paramètres géométriques

γ	Angle azimutal du capteur	Degrés [°]
γ_s	Angle azimutal solaire	Degrés [°]
δ	Déclinaison solaire	Degrés [°]
Φ	Latitude	Degrés [°]
α_s	Angle d'altitude solaire	Degrés [°]
θ	Angle d'incidence des rayons solaires	Degrés [°]
θ_i	Angle zénithal solaire	Degrés [°]

Rayonnements solaires

I	Rayonnement direct	Watt par mètre carré [W/m ²]
I_x	Rayonnement direct normal (perpendiculaire)	Watt par mètre carré [W/m ²]
D	Rayonnement diffus	Watt par mètre carré [W/m ²]
R	Rayonnement réfléchi	Watt par mètre carré [W/m ²]
G	Rayonnement global (horizontal)	Watt par mètre carré [W/m ²]
G_t	Rayonnement global sur une surface inclinée	Watt par mètre carré [W/m ²]
G_h	Irradiation globale horizontale	Watt-heure par mètre carré [Wh/m ²]

Constantes et indices

E_{sc}	Constante solaire	≈ 1361 [W/m ²]
K_t	Indice de clarté	Sans unité

ε_o	Facteur de correction d'excentricité orbitale	Sans unité
-----------------	---	------------

Paramètres temporels et spatiaux

S_a	Durée moyenne quotidienne d'insolation	Heures [h]
-------	--	------------

S_o	Durée maximale théorique d'insolation	Heures [h]
-------	---------------------------------------	------------

ρ	Albédo	Sans unité (0 à 1)
--------	--------	--------------------

Variables supplémentaires

E	Équation du temps	Minutes [min]
-----	-------------------	---------------

Y	Année	[-]
-----	-------	-----

X_{norm}	Données normalisées	Dépend du contexte
------------	---------------------	--------------------

Nomenclature

Liste des abréviation

Acronyme	Nom complet (Français)	Domaine d'application
ANN	Réseau de Neurones Artificiels	Apprentissage automatique, modélisation de systèmes complexes
ANFIS	Système d'Inférence Neuro-Flou Adaptatif	Systèmes hybrides (logique floue + réseaux de neurones), contrôle intelligent
GA	Algorithme Génétique	Optimisation, recherche de solutions dans des espaces complexes
PSO	Optimisation par Essaim de Particules	Optimisation collective, métaheuristique inspirée du comportement des essaims
EEMD	Décomposition Empirique Modale d'Ensemble	Traitement du signal non linéaire et non stationnaire
LSTM	Mémoire à Long-Court Terme (Long Short-Term Memory)	Réseaux de neurones récurrents, prédiction de séries temporelles
RMSE	Erreur Quadratique Moyenne Racine	Métrique d'évaluation de la précision des modèles (erreur moyenne)
R	Coefficient de Détermination (R^2)	Statistique, mesure de la qualité d'ajustement d'un modèle
MAE	Erreur Absolue Moyenne	Métrique d'évaluation de la précis



Introduction Générale



Introduction générale

L'énergie solaire, en tant qu'énergie renouvelable et disponible en grande quantité, est au cœur de la transition énergétique à l'échelle mondiale. Elle constitue une option pérenne aux énergies fossiles, aidant ainsi à diminuer les émissions de gaz à effet de serre et à combattre le réchauffement climatique. Toutefois, une utilisation optimale de cette ressource requiert une évaluation exacte du rayonnement solaire, spécialement dans les régions sèches où le niveau d'ensoleillement est élevé mais où les conditions climatiques extrêmes représentent des enjeux techniques et scientifiques considérables.

Les régions désertiques, comme le Sahara, offrent un potentiel solaire exceptionnel, mais leur environnement climatique unique caractérisé par des températures extrêmes, une faible humidité et des phénomènes atmosphériques spécifiques rend la prédiction du rayonnement solaire particulièrement complexe. Une estimation précise est pourtant cruciale pour optimiser la conception et la gestion des systèmes photovoltaïques, améliorer l'efficacité des centrales solaires, et soutenir des secteurs tels que l'agriculture et la météorologie dans ces régions.

Les modèles traditionnels de prédiction du rayonnement solaire, qu'ils soient empiriques, physiques ou statistiques, ont montré des limites dans leur capacité à capturer les relations complexes entre les variables météorologiques et le rayonnement solaire dans des environnements extrêmes. Ces dernières années, les techniques d'apprentissage automatique, notamment les réseaux de neurones artificiels (ANN) et les systèmes d'inférence neuro-flous adaptatifs (ANFIS), ont émergé comme des outils prometteurs pour surmonter ces limitations. Cependant, l'optimisation des paramètres de ces modèles reste un défi, nécessitant l'intégration de méthodes d'optimisation avancées telles que les algorithmes méta-heuristiques.

Cette thèse propose une approche innovante pour la prédiction du rayonnement solaire global dans les régions arides, en combinant les avantages des modèles d'apprentissage automatique (ANN et ANFIS) avec des algorithmes méta-heuristiques (algorithme génétique - GA et optimisation par essaim de particules - PSO). L'objectif est de développer des modèles hybrides capables de surmonter les limitations des méthodes traditionnelles, en améliorant à la fois la précision et la robustesse des prédictions dans des conditions climatiques extrêmes.

Les travaux de cette thèse s'articulent autour de plusieurs axes :

- La collecte et le prétraitement des données météorologiques dans trois régions arides du sud de l'Algérie (Adrar, Tindouf et Tamanrasset), en tenant compte des spécificités climatiques locales.
- Le développement et l'optimisation des modèles hybrides, en intégrant des techniques d'apprentissage automatique et des algorithmes méta-heuristiques pour affiner les paramètres des modèles.
- L'évaluation des performances des modèles à l'aide de métriques statistiques rigoureuses, en comparant les résultats avec ceux des modèles traditionnels.
- L'application des modèles développés à des problématiques pratiques, telles que l'estimation du potentiel solaire, l'optimisation de la gestion de l'eau et de l'énergie, et l'amélioration des prévisions météorologiques.

Les résultats obtenus démontrent que les modèles hybrides, en particulier ceux combinant ANN et PSO, surpassent les modèles traditionnels en termes de précision et de robustesse. Ces avancées ouvrent de nouvelles perspectives pour l'exploitation de l'énergie solaire dans les régions arides, tout en offrant des outils précieux pour les décideurs et les ingénieurs.

Cette thèse est organisée en cinq chapitres, chacun abordant des aspects spécifiques de la problématique et des méthodologies développées :

Le premier chapitre présente une revue approfondie des concepts fondamentaux liés à l'énergie solaire, au rayonnement solaire et aux techniques de modélisation. Il aborde également les défis spécifiques posés par les régions arides et les limites des modèles traditionnels. Ce chapitre sert de base théorique pour les travaux ultérieurs.

Le deuxième chapitre se concentre sur les méthodes d'estimation du rayonnement solaire, en distinguant les modèles empiriques, physiques et basés sur l'intelligence artificielle. Il explore également les applications de ces méthodes dans des domaines tels que l'énergie solaire, l'agriculture et la météorologie, en mettant en lumière les défis spécifiques aux régions arides.

Le troisième chapitre décrit la méthodologie adoptée pour la collecte, le prétraitement et l'analyse des données météorologiques et satellitaires. Il présente également les outils informatiques et logiciels utilisés pour le développement des modèles, ainsi que les techniques d'optimisation par méta-heuristiques.

Le quatrième chapitre présente les résultats obtenus à partir des modèles développés (ANN-GA; ANN-PSO; ANFIS-GA et ANFIS-PSO), en mettant l'accent sur leurs performances comparatives. Il analyse l'impact des données d'entrée sur la précision des modèles et explore leurs applications pratiques, notamment dans l'estimation du potentiel solaire et l'optimisation de la gestion de l'énergie.

Le cinquième chapitre présente en détail la méthode hybride innovante (EEMD-GA-LSTM) proposée dans cette thèse. Il explique les principes de la décomposition par EEMD, de l'optimisation par algorithmes génétiques et de la prédiction via LSTM. La méthodologie de mise en œuvre, incluant le prétraitement des données, la configuration des modèles et la sélection des hyperparamètres, est décrite. Enfin, les performances du modèle hybride sont évaluées à travers une validation croisée et comparées à celles des modèles existants.

La thèse se conclut par une synthèse des principaux résultats et contributions, en mettant en avant les avancées réalisées grâce à l'approche hybride proposée. Les limites de l'étude sont également discutées, et des pistes de recherche futures sont suggérées pour approfondir les travaux et étendre les applications de cette méthode à d'autres régions ou domaines.



Chapitre I :

Énergie Solaire :

Origine, Mesure

et Influences



Introduction

L'exploitation de l'énergie solaire et l'optimisation de sa collecte à travers les systèmes photovoltaïques reposent sur une distribution précise de l'irradiation solaire. Celle-ci est influencée par des paramètres géographiques tels que la latitude, la longitude et l'altitude, ainsi que par des facteurs météorologiques comme la nébulosité et l'angle d'incidence des rayons solaires selon la saison. Par ailleurs, des variables astronomiques, notamment la position du soleil dans le ciel au fil de la journée et de l'année, déterminent la quantité d'énergie solaire disponible sur le site d'installation du système photovoltaïque.

Dans ce chapitre, nous approfondissons ces notions afin d'établir un cadre complet. Nous débutons par les principes fondamentaux de l'énergie solaire, expliquant son transfert du soleil à la Terre et les méthodes de mesure utilisées. Ensuite, nous analysons les paramètres influençant l'intensité du rayonnement solaire atteignant le sol et leur impact sur l'efficacité des systèmes photovoltaïques. Nous introduisons également l'effet photovoltaïque et décrivons le fonctionnement des cellules solaires, essentielles à la conversion de la lumière en électricité.

Enfin, nous passons en revue les principaux composants d'un système photovoltaïque, depuis les panneaux solaires jusqu'aux dispositifs de stockage et de conversion de l'énergie, en mettant en lumière leur rôle dans la conception et le bon fonctionnement de ces systèmes innovants.

1. Le soleil

Le Soleil, situé à environ 150 millions de kilomètres de la Terre, constitue la principale source d'énergie pour notre planète. Il émet un rayonnement électromagnétique de manière uniforme, que l'on peut assimiler, en première approximation, à celui d'un corps noir à une température d'environ 5800 K. Ce rayonnement se propage dans l'espace sous forme d'ondes électromagnétiques et atteint la Terre en près de huit minutes après son émission. La majeure partie de cette énergie rayonnée, soit environ 98 %, se situe dans une plage de longueurs d'onde comprise entre 0,25 et 3 μm , couvrant ainsi les domaines de l'ultraviolet, du visible et de l'infrarouge proche [1].

Le Soleil est une étoile de forme pseudo-sphérique, semblable à une immense boule de gaz incandescent. Il est principalement composé d'hydrogène (80 %) et d'hélium (19 %), le 1 % restant étant un mélange de plus d'une centaine d'éléments, incluant pratiquement tous ceux présents sur Terre [2].

Bien qu'il soit considéré comme une étoile de taille moyenne, le Soleil représente à lui seul 99 % de la masse totale du système solaire. Sa structure interne n'est pas homogène et se divise en plusieurs zones. Son noyau, qui concentre 40 % de sa masse, est le siège des réactions thermonucléaires produisant

90 % de son énergie sous forme de rayonnements gamma et X. Cette région s'étend sur environ 25 000 km et présente des conditions extrêmes [1.3] :

- Température avoisinant 15 millions de degrés Celsius ;
- Densité d'environ 10^5 kg/m^3 ;
- Pression atteignant 10^9 atmosphères.

Une partie de cette énergie rayonnée parvient jusqu'à la Terre. L'intensité énergétique reçue à la surface terrestre est appelée éclairement énergétique ou simplement éclairement. La moyenne annuelle de cet éclairement, mesurée sur un plan perpendiculaire aux rayons du Soleil et situé au sommet de l'atmosphère terrestre, est connue sous le nom de constante solaire. La figure 1 illustre la répartition spectrale du rayonnement solaire atteignant notre planète [4].

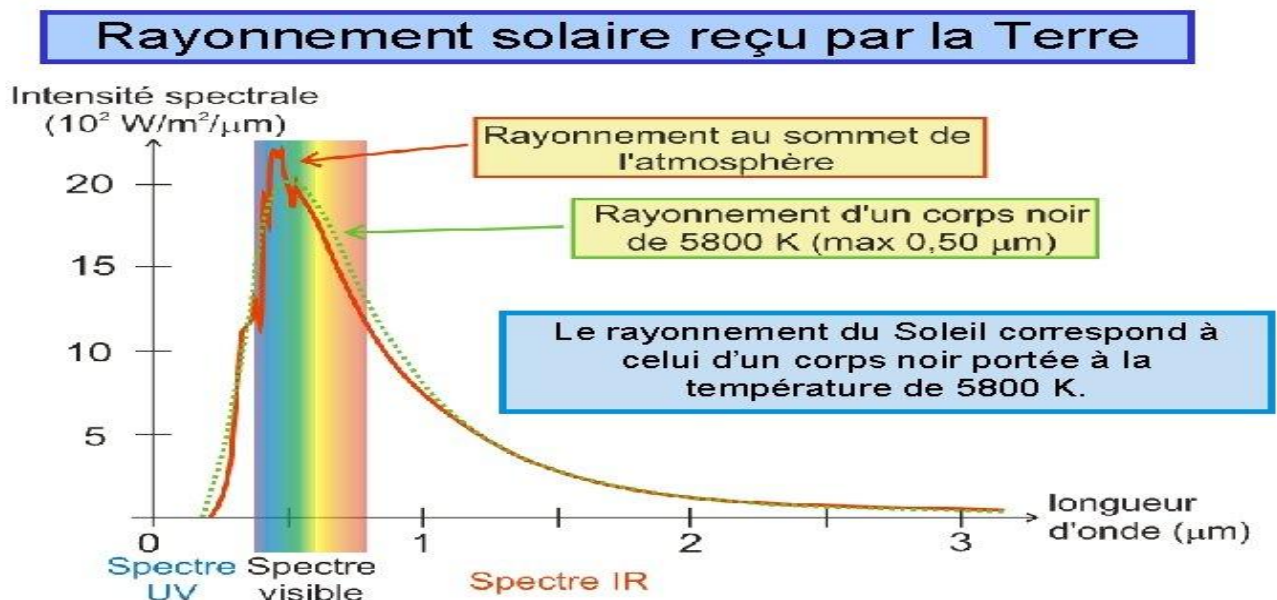


Figure1 : la distribution spectrale du rayonnement solaire reçu par la terre.[7]

2. Le système terre Soleil

La Terre, à l'instar des autres planètes de notre système solaire, orbite autour du Soleil dans un plan désigné sous le nom de plan de l'écliptique. Son axe de rotation présente une inclinaison de $23,5^\circ$ par rapport à la verticale de ce plan, influençant considérablement le cycle saisonnier. La Terre tourne entièrement sur son axe en 24 heures et orbite autour du Soleil en approximativement 365,25 jour.

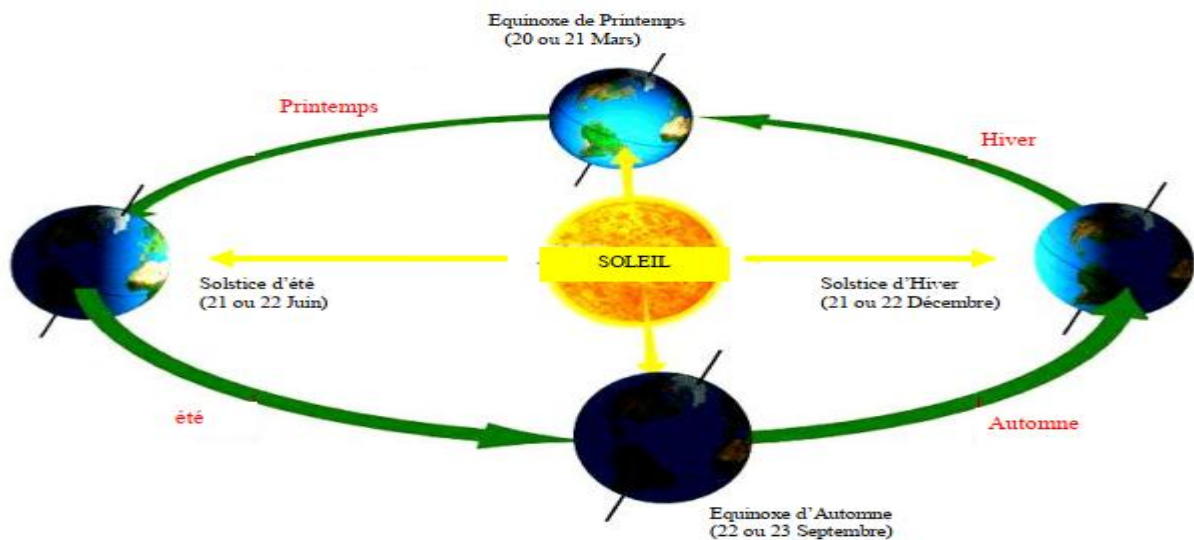


Figure 2 : Mouvement du globe terre par rapport au soleil.[8]

La position relative de la Terre par rapport au Soleil détermine la succession des saisons. Un paramètre clé dans ce phénomène est la déclinaison solaire, qui correspond à l'angle entre les rayons du Soleil et le plan équatorial terrestre. Cette déclinaison varie au cours de l'année : elle atteint $+23^\circ 27'$ lors du solstice d'été (vers le 21-22 juin), $-23^\circ 27'$ au solstice d'hiver (21-22 décembre), et s'annule lors des équinoxes (20-21 mars et 22-23 septembre), moments où le jour et la nuit ont une durée égale. Cette variation est à l'origine des saisons, et la déclinaison suit un cycle compris entre $-23^\circ 27'$ et $+23^\circ 27'$.

Par ailleurs, le mouvement de la Terre autour du Soleil n'est pas uniforme. Sa vitesse orbitale varie : elle est plus rapide en hiver et plus lente en été. Cette différence provoque une inégalité dans la durée des saisons. Par exemple, l'été dure environ 93,6 jours, le printemps 92,8 jours, l'automne 89,8 jours, et l'hiver seulement 89,0 jours. Cette variation découle directement de la deuxième loi de Kepler, aussi appelée loi des aires, qui stipule que la ligne reliant une planète au Soleil balaie des aires égales en des temps égaux.

Cela signifie que lorsque la Terre est au périhélie (le point de son orbite le plus proche du Soleil), sa vitesse est maximale, tandis qu'elle ralentit lorsqu'elle se trouve à l'aphélie (le point le plus éloigné).

Cette dynamique implique que la force gravitationnelle exercée par le Soleil agit constamment en direction de celui-ci, maintenant la Terre dans son orbite. [6]

3. Le bilan radiatif terrestre

Pour estimer l'énergie solaire reçue sur un site donné, il est essentiel de connaître sa position géographique ainsi que de comprendre les mouvements de la Terre autour du Soleil. Ces éléments influencent directement la quantité d'énergie solaire disponible à la surface terrestre.

3.1. Paramètres géographiques

La Terre est divisée en deux hémisphères par l'équateur :

- L'hémisphère Nord, situé au-dessus de l'équateur, en direction du pôle Nord
- L'hémisphère Sud, en dessous de l'équateur, vers le pôle Sud

Elle est également partagée en hémisphères Est et Ouest par le méridien de Greenwich, situé près de Londres, qui sert de référence pour mesurer les longitudes. [7]

3.1.1. La latitude

La latitude d'un point est une évaluation angulaire qui dénote son emplacement nord-sud en relation avec l'équateur. Elle est formulée en degrés : 0° à l'équateur, $+90^\circ$ au pôle Nord et -90° au pôle Sud. Dans l'hémisphère Nord, la latitude est exprimée en valeurs positives tandis que dans l'hémisphère Sud, elle est indiquée en valeurs négatives. On l'appelle généralement θ . [5]

3.1.2. La longitude

La longitude exprime la position est-ouest d'un lieu par rapport au méridien d'origine (Greenwich). C'est une mesure angulaire allant de : -180° à $+180^\circ$, ou de 180° Ouest à 180° Est. Notée φ , la longitude est également liée au temps : la Terre tournant de 360° en 24 heures, chaque degré de longitude correspond à 4 minutes de différence horaire. [6]

3.1.3. L'altitude

L'altitude d'un point représente sa hauteur par rapport à un niveau de référence, généralement le niveau moyen de la mer. Elle se mesure en mètres et permet de distinguer les zones de haute ou basse élévation. L'altitude influe également sur l'intensité du rayonnement solaire reçu, l'atmosphère étant plus mince en altitude.

Elle est souvent déterminée à l'aide de méthodes géodésiques ou topographiques telles que la triangulation. [5]

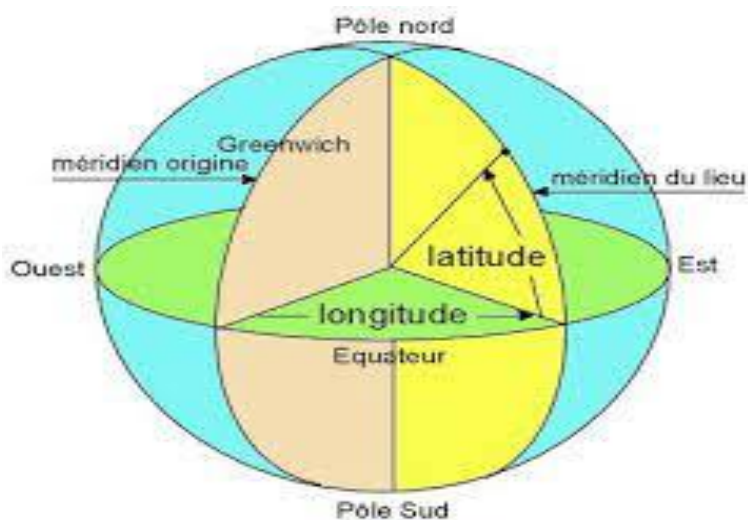


Figure 3 : Coordonnées géographiques du lieu.

3.2. Paramètres astronomiques

3.2.1. Repérage du soleil

On peut déterminer l'emplacement du Soleil dans le ciel à n'importe quel moment de la journée ou de l'année en utilisant deux systèmes différents de coordonnées. Ces dispositifs facilitent la détermination précise de la position visible du Soleil par rapport à un emplacement spécifique sur Terre, en considérant simultanément le déplacement de la Terre autour du Soleil et sa propre rotation.

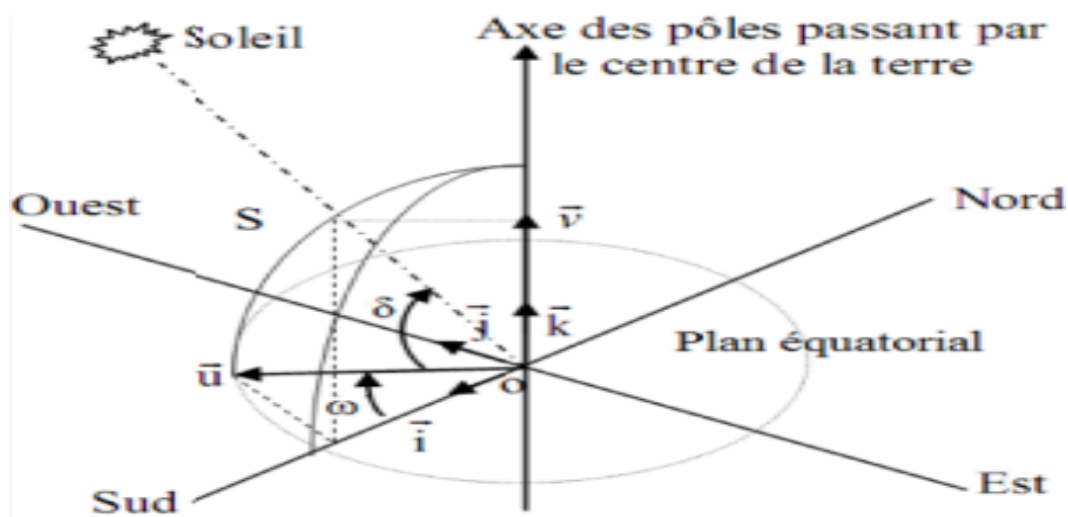


Figure 4 : Repère équatorial.

3.2.2. Déclinaison du soleil

La déclinaison du Soleil est l'angle formé entre la direction des rayons solaires et leur projection sur le plan équatorial terrestre. Cet angle varie tout au long de l'année en raison de l'inclinaison de l'axe terrestre et du mouvement de la Terre autour du Soleil. La déclinaison atteint :

- $+23^{\circ}27'$ lors du solstice d'été (21-22 juin),
- $-23^{\circ}27'$ lors du solstice d'hiver (21-22 décembre).

Divers modèles mathématiques ont été proposés pour calculer cette déclinaison, mais la formule la plus couramment utilisée est celle de Cooper : [9]

$$\delta = 23.45^{\circ} \sin \left| \frac{360}{365} (nj + 284)^{\circ} \right| \quad (1)$$

Où nj est le numéro du jour de l'année.

3.2.3. Angle horaire du soleil

L'angle horaire est une manière angulaire d'exprimer le temps solaire à un instant donné. Plutôt que d'utiliser les heures classiques, on le mesure en degrés : 360° correspondent à 24 heures, soit 15° par heure. L'angle horaire ω représente l'angle entre le méridien local (passant par le Sud) et la projection de la direction du Soleil sur le plan équatorial. Il permet de suivre la course apparente du Soleil dans le ciel. Il est défini par la formule : [10]

$$\omega = 15(12 - TSV) \text{ En degré} \quad (2)$$

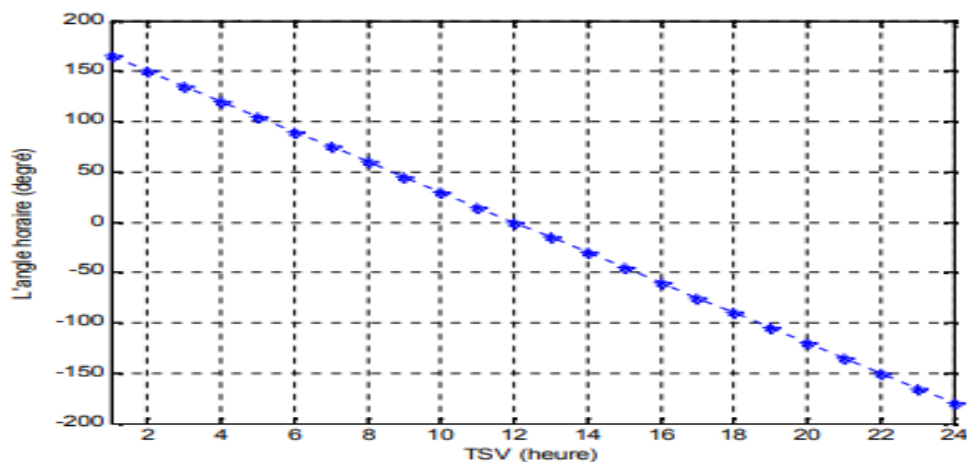


Figure 5 : Variation de l'angle horaire en fonction du temps.

3.3.Coordonnées horizontales

Le plan de l'horizon astronomique et la direction verticale du lieu constituent le repère horizontal. Dans ce cadre, on établit la définition des coordonnées de la manière suivante :[11-12]

- la hauteur du soleil, 'h'
- l'azimut 'az'

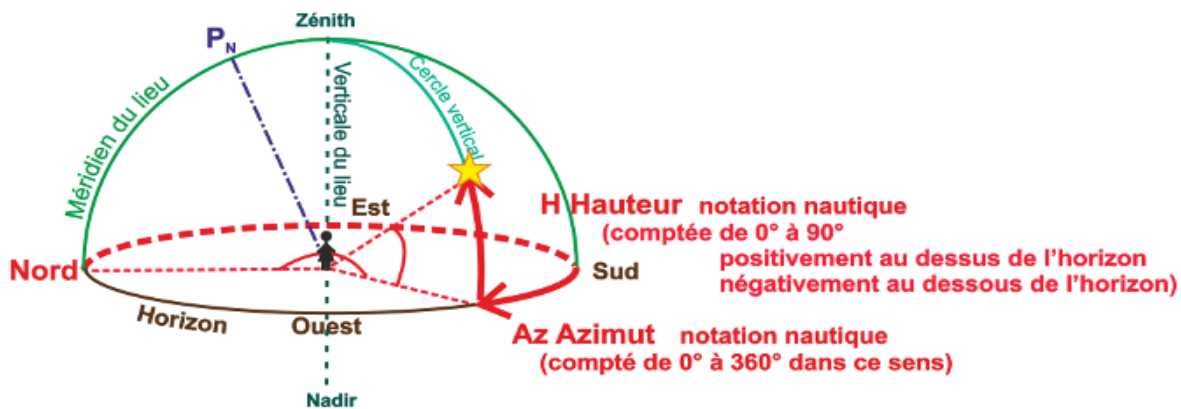


Figure 6: cordonnées horizontales.

3.3.1. La hauteur du soleil

La hauteur représente l'angle créé entre la trajectoire du Soleil et sa projection sur le plan horizontal. Elle est exactement nulle lors du lever et du coucher astronomiques du Soleil, atteignant son niveau le plus élevé à midi, en temps solaire authentique. La formule suivante fournit l'expression de la hauteur solaire:

$$\sin(h) = \cos(\delta) \cos(\theta) \cos(\omega) + \sin(\theta) \sin(\delta) \quad (3)$$

Avec :

δ : La déclinaison

θ : La latitude de lieu

ω : L'angle horaire.

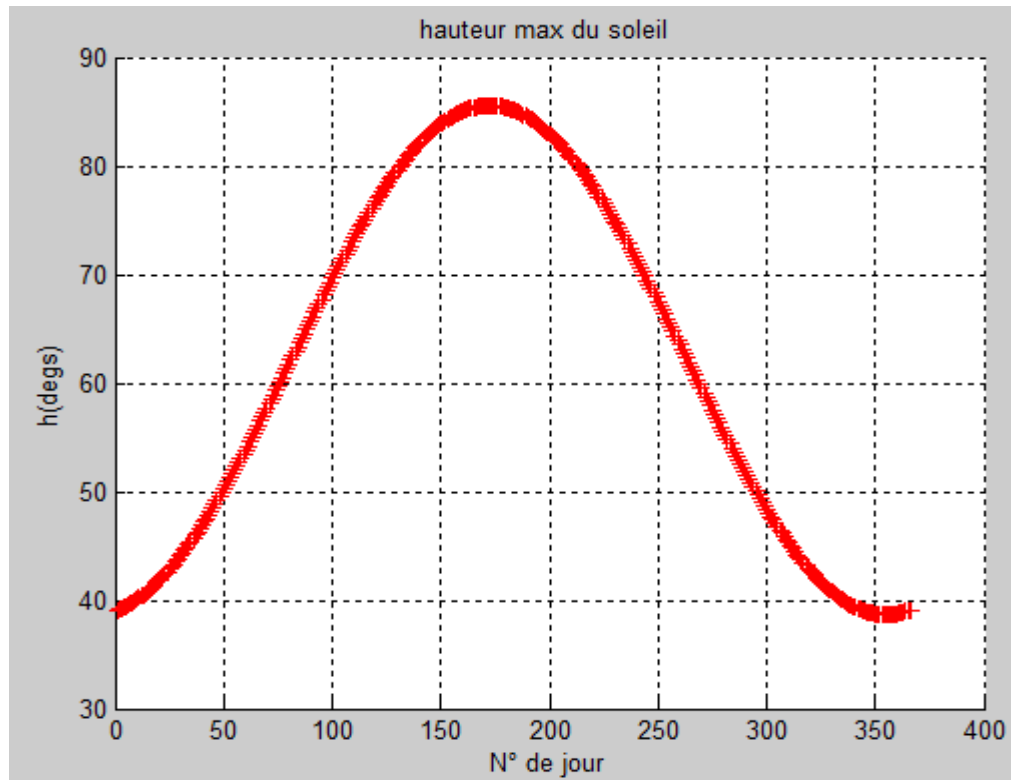


Figure 7 : Variation de la hauteur du soleil au cours d'une année.

3.3.2. Azimut du soleil

L'azimut du Soleil est l'angle mesuré entre la projection de la direction du Soleil sur le plan horizontal et la direction du sud. Il est donné par la relation suivante :[13]

$$\sin(a_z) = \frac{\cos(\delta) \sin(\omega)}{\cos(h)} \quad (4)$$

4. Les temps solaires

Pour les applications liées à l'énergie solaire, il est nécessaire d'utiliser le temps solaire vrai, calculé en tenant compte de l'écart entre le temps solaire moyen et le temps local.

4.1. Le temps solaire moyen

La rotation de la Terre sur elle-même introduit la notion de temps solaire moyen (TSM). Il est défini comme le temps séparant, en moyenne, deux passages successifs du Soleil au méridien d'un lieu donné. Une rotation complète de la Terre prend 24 heures.[14]

4.2. Le temps universel

Le Temps Universel (TU) est défini comme l'heure du passage du Soleil moyen au méridien d'origine (Greenwich). À la longitude zéro, le temps solaire moyen (TSM) coïncide donc avec le TU. La différence entre le TSM local et le TU est appelée correction de longitude. La relation entre les deux s'écrit ainsi:

$$TSM = TU \pm L/15 \quad (5)$$

L : longitude du lieu, (+) pour longitude Est, (-) pour longitude Ouest.

4.3. Le temps légal

Le temps légal (ou local) est le temps officiel du site, il diffère par rapport au temps mondial de Greenwich par un décalage (+ ou -) exprimé en heures, est exprimé sous la forme :

$$TL = TU + \Delta H \quad (6)$$

En tenant compte ΔH étant le décalage horaire entre le méridien de Greenwich et l'Etat considéré, ($\Delta H=1$ heure pour l'Algérie).[11]

4.4. Le temps solaire vrai

Le temps solaire vrai (TSV) est le temps mesuré entre deux passages consécutifs du Soleil au méridien d'un lieu. Il diffère du temps solaire moyen. Cette différence est appelée équation du temps (ET) et varie selon le jour de l'année :

$$TSV = TSM + ET \quad (7)$$

4.5. L'équation du temps

L'équation du temps prend en compte la variation de la vitesse de rotation de la Terre et est exprimée par la relation suivante:

$$ET = 9.9 \sin[2(0.986 j + 100)] - 7.7 \sin (0.986 j - 2) \quad (8)$$

Ainsi, un jour solaire peut varier entre 24 h 00 min 30 s et 23 h 59 min 40 s. D'un jour à l'autre, l'écart est de quelques secondes (jusqu'à 30 secondes vers fin décembre). Ces variations accumulées entraînent un décalage progressif de l'instant du passage du Soleil au méridien (midi solaire) par rapport au midi donné par une horloge régulière, pouvant atteindre jusqu'à 17 minutes sur l'année.

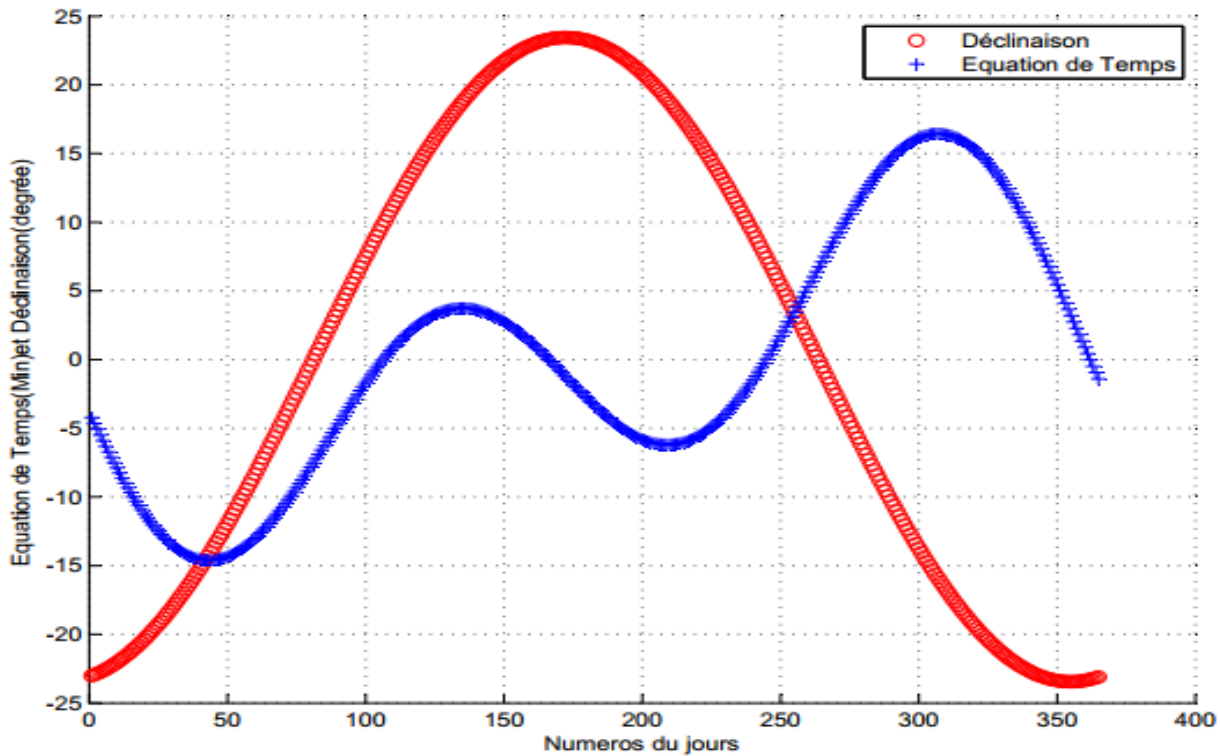


Figure 8 : Equation du temps et déclinaison en fonction du numéro de jour de l'année.

4.6. Transition du temps légal vers le temps solaire vrai

En général, pour convertir le temps légal local (TL) en temps solaire vrai (TSV), on utilise l'expression suivante :

$$TSV = TL - \Delta H + \left(\frac{ET + 4\varphi}{60} \right) \quad (9)$$

ΔH : est le décalage horaire par rapport au méridien de Greenwich. (Égale 1 pour l'Algérie).

TL: est le temps légal ou temps local standard, donné par une montre.

φ : longitude du lieu.

5. Émission du soleil :

5.1. Nature du rayonnement solaire :

Les ondes électromagnétiques composent le rayonnement solaire. On désigne par rayonnement électromagnétique (REM) un phénomène où il agit comme un champ de forces dont les fluctuations influencent les caractéristiques électriques et magnétiques des substances. La lumière représente l'élément visible du REM, que l'œil humain est capable de percevoir.

La longueur d'onde désigne la distance qui sépare deux points identiques successifs d'une onde, alors que la fréquence définit le nombre de cycles qui traversent un point fixe en une seconde. Plus la fréquence augmente, plus la longueur d'onde diminue. Ces fluctuations voyagent dans le vide à la vitesse de la lumière, soit environ 300 000 km/s.

Propagation des ondes électromagnétiques :

Le rayonnement solaire atteint la Terre en environ 499 secondes, soit 8 minutes et 19 secondes. D'après la théorie corpusculaire de Max Planck, tout rayonnement à une fréquence f peut être envisagé comme un ensemble de photons dont l'énergie individuelle est liée à cette fréquence :

$$E = h\nu \quad (10)$$

Ainsi, les rayonnements de courte longueur d'onde (haute fréquence) sont plus énergétiques.

Notre œil ne perçoit qu'une partie du rayonnement solaire, appelée domaine visible, comprise entre 0,40 μm et 0,70 μm (certains auteurs étendent cette limite à 0,77 μm pour équilibrer les contributions énergétiques du visible et du proche infrarouge).

Le Soleil émet cependant sur une large gamme de longueurs d'onde, allant :

- des rayons gamma (longueurs d'onde $< 10^{-12}$ m),
- aux ondes radioélectriques (longueurs d'onde jusqu'à 1000 m), en passant par les rayons X, ultraviolets, visible, infrarouge et hyperfréquences.

La majorité de l'énergie solaire (99,2 %) hors atmosphère se situe entre 0,2 μm et 4 μm . Après absorption par la vapeur d'eau atmosphérique, le spectre au sol est limité à environ 2,5 μm (Figure 1.).[15]

5.2. Propriétés des ondes électromagnétiques

Les propriétés fondamentales des ondes électromagnétiques, essentielles à la compréhension de l'énergie solaire, sont : [8]

- ❖ La réflexion
- ❖ L'absorption
- ❖ La diffusion.

5.3. Spectre du soleil

Le Soleil émet un rayonnement électromagnétique principalement dans le domaine de la lumière, composé de : [3]

- La lumière visible dont les longueurs d'onde vont de $0.4\mu\text{m}$ et $0.8\mu\text{m}$, elle représente 48% du rayonnement total.
- Le rayonnement ultraviolet (UV) à des longueurs d'onde inférieures à $0.4\mu\text{m}$ constitue 6% de l'ensemble du rayonnement.
- Le rayonnement infrarouge (IR) à des longueurs d'onde supérieures à $0.8\mu\text{m}$ représente 46% du rayonnement total.

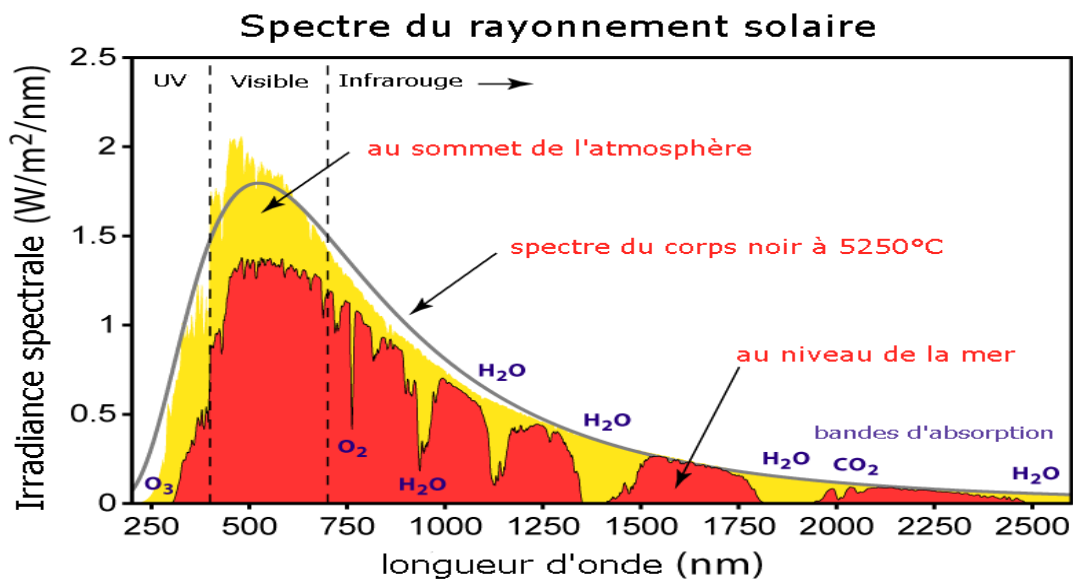


Figure 9 : Spectre de rayonnement solaire [3].

6. Le rayonnement solaire hors atmosphère

Le rayonnement solaire extraterrestre (hors atmosphère) correspond au rayonnement solaire qui atteint la limite supérieure de l'atmosphère terrestre.

6.1.Le couple Terre-soleil

Le couple Terre-Soleil est fondamental pour de nombreuses dynamiques et phénomènes influençant notre planète et notre vie quotidienne. À travers différentes dimensions, cette interaction complexe façonne l'environnement terrestre.

6.1.1. Orbite et Saisons:

La Terre décrit une orbite elliptique autour du Soleil. Cette orbite détermine l'alternance des saisons : printemps, été, automne et hiver. L'inclinaison de l'axe de rotation terrestre joue un rôle déterminant dans ce phénomène, influençant directement la quantité de lumière solaire reçue aux différentes latitudes au cours de l'année.

6.1.2. Lumière et Énergie:

Le Soleil constitue la principale source de lumière et d'énergie pour la Terre. Les photons émis par le Soleil traversent l'espace pour atteindre notre planète, où ils sont absorbés par l'atmosphère, les océans et la surface terrestre. Cette énergie est essentielle à la vie, alimentant notamment la photosynthèse, régulant le climat et constituant la base de nombreuses énergies renouvelables, telles que l'énergie solaire photovoltaïque.

6.2.Climat et Météorologie:

L'interaction entre la Terre et le Soleil conditionne le climat et la météorologie. Les variations d'ensoleillement selon les latitudes génèrent des courants atmosphériques et océaniques, influençant ainsi les modèles de circulation atmosphérique, les précipitations et les températures à l'échelle mondiale.

6.3.Espace et Exploration:

L'étude du couple Terre-Soleil est également essentielle pour la recherche spatiale et l'exploration. Les missions spatiales, qu'elles soient destinées à Mercure ou à Mars, permettent de mieux comprendre l'histoire de notre système solaire et les effets des phénomènes solaires sur les technologies terrestres.

En résumé, le couple Terre-Soleil est une relation dynamique et indispensable, qui modèle notre monde de multiples façons, influençant à la fois notre climat, notre météo et notre compréhension de l'univers.

6.4.Correction de la distance Terre soleil

Les fluctuations annuelles de la distance entre la Terre et le Soleil résultent de la nature elliptique de l'orbite terrestre. Elles provoquent des variations du rayonnement solaire hors de l'atmosphère (éclairage solaire direct extraterrestre). La valeur de correction associée à cette modification de distance

est calculée en fonction de la localisation de la Terre sur son orbite, en fonction du jour de l'année, conformément à l'équation suivante :

$$E_0 = \left(r_0/r\right)^2 = 1 + 0.033 \cos \left[\left(\frac{2\pi nj}{365}\right)\right] \quad (11)$$

Avec :

r_0 : la distance moyenne entre la Terre et le soleil (150 millions Km).

r : la distance actuelle entre la Terre et le soleil.

6.5. La constante solaire

La constante solaire est la quantité d'énergie solaire reçue par une surface qui se trouve à angle droit par rapport aux rayons solaires, placée à la limite supérieure de l'atmosphère, environ à 150 millions de km. Elle est autonome par rapport aux conditions climatiques, mais elle est liée à la distance entre la Terre et le Soleil. Elle peut être formulée comme suit :

$$I_0 = \sigma T_s^4 (R_s/D_{t,s})^2 \quad (12)$$

T_s : Température du soleil égale à 5800 K.

R_s : Rayon du soleil (696.106 m).

$D_{t,s}$: Distance moyenne Terre soleil.

σ : Constante de Stéphane Boltzmann = $5.67.10^{-8} \text{ W.m}^{-2} .\text{K}^{-4}$.

En moyenne, la Terre reçoit 1367 W/m^2 en haute atmosphère, à la verticale face au Soleil. La constante solaire représente donc la puissance solaire reçue, en l'absence d'atmosphère, par unité de surface perpendiculaire aux rayons solaires.

6.6. Correction de la constante solaire

La constante solaire corrigée s'exprime par :

$$I_{sc}(nj) = I_0 E_0 \quad (13)$$

I_{sc} : est la constante solaire corrigée.

I_0 : est la constante solaire moyenne qui est égale à 1367 W.m^{-2} .

La figure ci-dessous illustre l'évolution annuelle de la constante solaire :

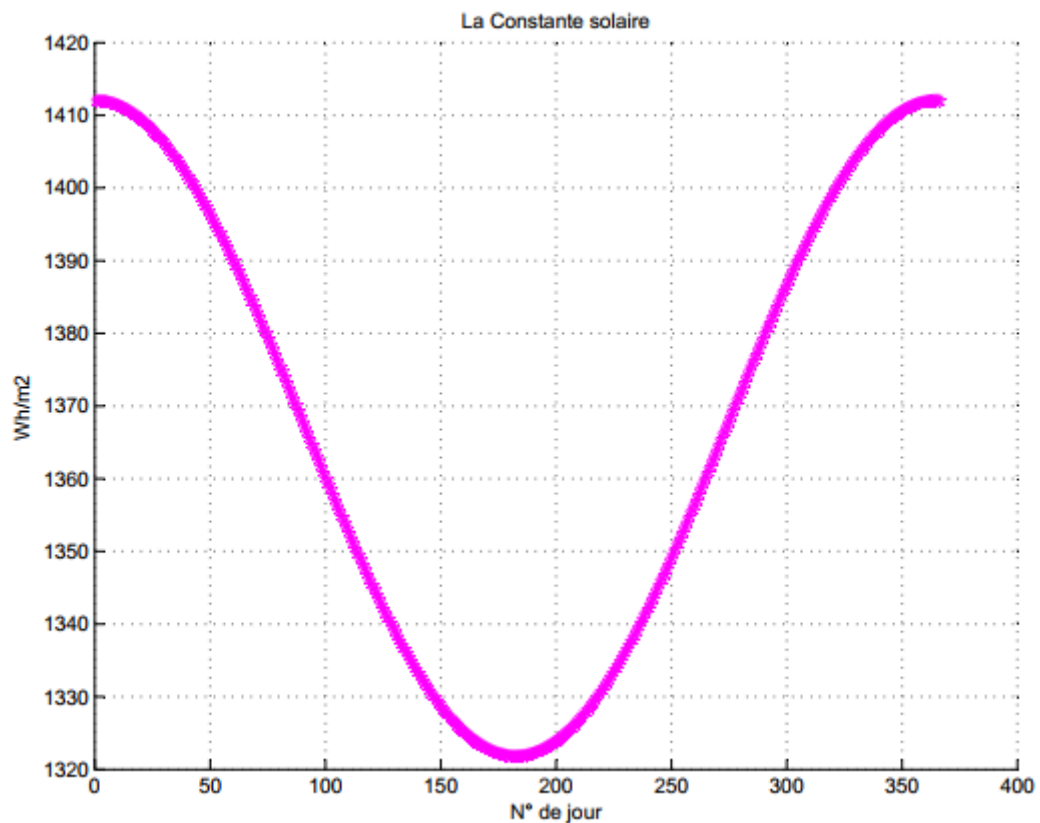


Figure 10 : Évolution annuelle de la constante solaire.

On observe un maximum en janvier avec 1413 W/m² et un minimum au début de juillet avec 1320 W/m².

6.7. Flux solaire sur plan horizontal

Pour un emplacement spécifique, l'intensité du rayonnement solaire reçu sur un plan horizontal hors de l'atmosphère est déterminée par la projection du flux normal sur ce plan horizontal:[4]

$$I_{0,h} = I_{SC} \sin(h) \quad (14)$$

6.8.Irradiation journalière sur un plan horizontal

L'irradiation globale journalière sur une surface horizontale, $G_0(nj)$, s'obtient par l'intégration du flux solaire entre le lever et le coucher du Soleil :[3]

$$G_0(nj) = \int_{t_l}^{t_c} I_{sc}(nj) \sin(h(TSV)) dt \quad (15)$$

où t_l et t_c sont respectivement les heures de lever et de coucher du Soleil. Après intégration, on obtient l'expression suivante pour le rayonnement extraterrestre quotidien sur une surface horizontale :

$$G_0(nj) = \frac{24}{\pi} I_{sc}(nj) [\cos(\varphi) \cos(\delta) \sin(\omega_s) + \omega_s \sin(\varphi) \sin(\delta)] \quad (16)$$

G_0 est une densité d'énergie exprimée en Wh/m².

Pour un plan incliné d'un angle β , l'irradiation globale journalière devient :

$$G_0(nj) = \frac{24}{\pi} I_{sc}(nj) [\cos(\varphi - \beta) \cos(\delta) \sin(\omega_s) + \omega_s \sin(\varphi - \beta) \sin(\delta)] \quad (17)$$

7. Mesures et Analyse du Rayonnement Solaire

L'analyse du rayonnement solaire est essentielle à la conception et à l'optimisation des systèmes photovoltaïques. Elle nécessite une collecte précise des données pour comprendre les variations du rayonnement et leur impact sur les systèmes énergétiques.

7.1.Techniques de Mesure du Rayonnement Solaire

Plusieurs instruments sont utilisés pour caractériser les différentes composantes du rayonnement solaire, chacun répondant à des besoins spécifiques :

7.1.1. Pyranomètres : Ces instruments mesurent le rayonnement global, c'est-à-dire la somme du rayonnement direct et diffus reçu par une surface horizontale. Ils sont équipés de dômes en verre ou en quartz, qui permettent de transmettre la lumière solaire tout en bloquant les longueurs d'onde infrarouges non pertinentes. Leur conception vise à assurer une mesure précise, indépendante de l'angle d'incidence du rayonnement [6].



Figure 11 : Pyranomètres

7.1.2. Pyrhéliomètres: Destinés à mesurer exclusivement le rayonnement direct, les pyrhéliomètres utilisent un tube de collimation pour limiter l'angle de vue à une étroite portion du ciel centrée sur le disque solaire. Le rayonnement est concentré sur un capteur thermique ou photovoltaïque, offrant ainsi une mesure précise du flux solaire direct normal [16].



Figure 12 : Pyrhéliomètres

7.1.3. Albédo-mètres : Composés généralement de deux pyranomètres montés dos à dos (l'un orienté vers le ciel, l'autre vers le sol), les albédo-mètres mesurent à la fois le rayonnement incident et réfléchi. Ils permettent de déterminer l'albédo, c'est-à-dire la fraction de l'énergie solaire réfléchie par une surface, un paramètre essentiel pour l'analyse énergétique des surfaces claires (neige, sable, etc.) [17].



Figure 13 : Albédo-mètre

7.1.4. Héliographes : Les héliographes sont utilisés pour enregistrer la durée quotidienne d'ensoleillement. Le modèle classique, tel que l'héliographe de Campbell-Stokes, utilise une sphère de verre pour concentrer les rayons solaires sur une bande de papier sensible à la chaleur. La longueur de la trace brûlée est directement proportionnelle à la durée d'ensoleillement.



Figure 14 : Héliographes

Ces instruments permettent de collecter des données précises sur les différentes composantes du rayonnement : direct, diffus, global et réfléchi, éléments indispensables pour la modélisation et l'optimisation des systèmes photovoltaïques.

7.2. Utilisation des Données Satellitaires

Dans les régions sahariennes, où les stations météorologiques au sol sont peu nombreuses, les satellites jouent un rôle crucial dans l'évaluation du potentiel solaire :

7.2.1. METEOSAT : Cette famille de satellites fournit des images haute résolution toutes les 15 minutes, permettant de suivre en temps quasi réel l'évolution de la couverture nuageuse et d'estimer le rayonnement solaire incident à la surface.

- 7.2.2. EUMETSAT:** Organisation spécialisée dans la télédétection météorologique et climatique, EUMETSAT exploite plusieurs satellites météorologiques fournissant des données précieuses pour la surveillance de l'irradiation solaire à large échelle géographique.
- 7.2.3. Hélio-Clim:** Base de données élaborée à partir d'images de satellites géostationnaires, HelioClim propose des estimations du rayonnement solaire à haute résolution spatiale et temporelle, très utilisées pour les études de dimensionnement photovoltaïque.

Les données satellitaires sont traitées à l'aide d'algorithmes complexes combinant des indices de nébulosité, les paramètres géographiques et les modèles radiatifs atmosphériques pour estimer précisément l'irradiation solaire au sol.

8. Impact des Conditions Atmosphériques

Les conditions atmosphériques modulent de manière importante le rayonnement solaire disponible :

- 8.1.Nébulosité:** Réduit considérablement le rayonnement direct mais peut parfois augmenter le rayonnement diffus par diffusion multiple (effet lumineux sous ciel couvert).
- 8.2.Pollution atmosphérique :** Les particules fines (PM2.5, PM10) absorbent et diffusent la lumière, diminuant la quantité de rayonnement atteignant le sol. Ceci est particulièrement critique dans les grandes villes ou les régions industrielles.
- 8.3.Vapeur d'eau :** La vapeur d'eau absorbe certaines longueurs d'onde du rayonnement solaire, notamment dans l'infrarouge. Cela impacte la quantité d'énergie thermique transmise.

Ces effets sont généralement modélisés à l'aide de données climatiques historiques ou de prévisions météorologiques en temps réel.

9. Bases de Données Climatiques

Pour des analyses précises, différentes bases de données climatiques sont utilisées :

- 9.1.Hélio-Clim:** Fournit des estimations journalières à haute résolution spatiale du rayonnement solaire en Europe, Afrique du Nord et Moyen-Orient. Les données sont issues d'images satellites et de modèles physiques.
- 9.2.NASA SSE (Surface Solar Energy):** Propose des données gratuites et globales sur l'énergie solaire, utiles pour des évaluations préliminaires de sites partout dans le monde.
- 9.3.PVGIS (Photovoltaic Geographical Information System):** Donne accès à des outils interactifs permettant d'estimer la production photovoltaïque annuelle ou mensuelle, en intégrant les pertes dues à la température, l'angle d'inclinaison, etc.

10. Modélisation et Simulation

Les données climatiques et géographiques sont ensuite utilisées pour :

- Identifier les périodes de forte production : Par exemple, pour planifier la maintenance pendant les périodes de faible production.
- Concevoir des systèmes adaptés : Choix de l'inclinaison optimale, dimensionnement des panneaux, intégration de systèmes de stockage ou de suivi solaire.
- Évaluer l'impact des variations climatiques : Simulation de l'effet des années particulièrement nuageuses ou de l'augmentation des aérosols liée aux feux de forêts.

Des logiciels comme PVSyst, HOMER, SAM (System Advisor Model) ou des approches machine learning sont couramment utilisés pour ces étapes.

Conclusion:

Ce chapitre a établi les bases théoriques et scientifiques indispensables à la compréhension des principes fondamentaux du rayonnement solaire et de son exploitation dans les régions sahariennes. Les concepts clés, notamment les paramètres géographiques, astronomiques et climatiques, ont été analysés en détail pour illustrer leur rôle déterminant dans l'estimation et l'optimisation de cette ressource naturelle.

La présentation des méthodes traditionnelles de mesure et de modélisation, ainsi que des outils modernes tels que les satellites et les bases de données climatiques, a mis en lumière l'évolution continue des approches visant à améliorer la précision des estimations. En particulier, les techniques avancées de mesure, comme les pyranomètres et les héliographes, ainsi que les systèmes satellitaires tels que METEOSAT et HelioClim, soulignent l'importance de disposer de données fiables et complètes dans des environnements complexes comme le Sahara.

Les défis liés à la variabilité spatio-temporelle du rayonnement solaire, associés aux conditions extrêmes des régions désertiques, renforcent la nécessité d'adopter des modèles robustes. Qu'ils soient empiriques, physiques ou fondés sur l'intelligence artificielle, ces modèles présentent chacun des avantages et des limites. Cependant, leur combinaison avec des approches innovantes, telles que les méta-heuristiques, ouvre des perspectives prometteuses pour surmonter les contraintes actuelles.

En conclusion, l'étude approfondie des mécanismes influençant le rayonnement solaire et des méthodologies d'estimation constitue une étape essentielle pour optimiser les systèmes photovoltaïques dans des régions à fort potentiel solaire. Ce cadre conceptuel servira de socle pour explorer, dans les chapitres suivants, des approches hybrides et des technologies avancées exploitant les dernières innovations en matière de modélisation et de simulation.

Reference:

- [1] Alfred H. pekarek, Geological perspectives of global climate change, Solar forcing of earth's climate. Minnesota, USA: Tulsa, 2001
- [2] M. Iqbal, 'An Introduction to Solar Radiation'. Academic Press, Department of Mechanical Engineering, University of British Columbia, Canada, 1983.
- [3] Pierre Bessemoulin et Jean Oliviéri, 'Le rayonnement solaire et sa composante ultraviolette', Météo-France, Service central d'exploitation de la météorologie Toulouse, La Météorologie 8eme série - n°31 - septembre 2000.
- [4] Blanc, P., & Wald, L. (2015). L'estimation du rayonnement solaire au sol par la nouvelle méthode Heliosat-4. *La Météorologie*, 90, 53-61.
- [5] Bouchouicha Kada, "MODELISATION MULTISPECTRALE DES IMAGES SATELLITAIRE - Application : Quantification du bilan d'énergie Sol-Atmosphère," USTO, Oran, Thèse de doctorat en sciences 2017.
- [6] BENATIALAH Djelloul, "Détermination du gisement solaire par imagerie satellitaire avec intégration dans un S.I.G pour le sud d'Algérie," Université Ahmed draya Adrar, Thèse de doctorat en sciences 2019.
- [7] Menguy G. et Schwartz M Benard R., Le rayonnement solaire, Conversion thermique et applications. Paris, France : Technique et Documentation, 1979.
- [8] Perrin de Brichambaut Vauge, Le gisement solaire - Evaluation de la ressource énergétique. paris, france: technique et documenation (loisier), 1982.
- [9] A.A.M Sayigh « SOLAR ENERGY », Solar Energy Engineering. Riadh, Saoudi Arabia: Academic press, INC, Orlando, Florida 32887, 1977.
- [10]. Bernard J., (2004), « Energie solaire calculs et optimisation », Ellipse Edition Marketing.
- [11] F. Yettou, A. Malek, M. Haddadi et A. Gama, Etude comparative de deux modèles de calcul du rayonnement solaire par ciel clair en Algérie, *Revue des Energies Renouvelables* vol. 12 N°2 pp.331 – 346, 2009
- [12] A. Moummi, N. Hamani, N.Moummi et Z. Mokhtari, 'Estimation du rayonnement solaire par deux approches semi empiriques dans le site de Biskra', (SPIE8)8ème Séminaire International sur la Physique Energétique, Algérie, Novembre 2006.

-
- [13] S. Benkaciali, K. Gairaa, Techniques de modélisations de l'irradiation solaire sur un plan incliné, *International Journal of Scientific Research & Engineering Technology (IJSET)*, Vol.3-pp.102-108, 2015.
- [14] BENCHACHA, M. (2020). *Développement d'un Outil Informatique Pour l'Estimation de l'Eclairement Solaire Incident Sur un Plan Horizontal en Algérie* (Doctoral dissertation, جامعة غرداية).
- [15] Cyril Voyant, (2011), « Prédiction de séries temporelles de rayonnement solaire global et de production d'énergie photovoltaïque à partir de réseaux de neurones artificiels », Thèse Doctorat en physique, Université de Corse-Pascal Paoli.
- [16] MacDonald, T. H. (1951). Some characteristics of the Eppley pyrliometer. *Monthly Weather Review*, 79(8), 153-159.
- [17] Wendisch, M., Müller, D., Schell, D., & Heintzenberg, J. (2001). An airborne spectral albedometer with active horizontal stabilization. *Journal of Atmospheric and Oceanic Technology*, 18(11), 1856-1866.



Chapitre II :

Estimation du

Rayonnement Solaire

dans les Régions Arides



Introduction

Le potentiel énergétique du soleil, ressource inépuisable et propre, suscite un intérêt croissant dans le monde entier. Le désert du Sahara, avec son ensoleillement exceptionnel, représente un gisement énergétique considérable qui pourrait révolutionner les systèmes énergétiques de nombreux pays. Néanmoins, exploiter pleinement cette ressource nécessite une connaissance précise du rayonnement solaire incident sur la surface terrestre.

L'estimation du rayonnement solaire dans des régions désertiques comme le Sahara se heurte à plusieurs défis. La variabilité spatio-temporelle du rayonnement, induite par des phénomènes météorologiques complexes tels que les tempêtes de sable, les changements de couverture nuageuse et les fluctuations de la transparence atmosphérique, rend difficile l'élaboration de modèles de prédiction fiables. De plus, le manque de données météorologiques complètes et continues dans ces régions, notamment en ce qui concerne les paramètres atmosphériques, constitue un obstacle majeur.

Les conséquences d'une estimation erronée du rayonnement solaire peuvent être importantes, tant sur le plan économique qu'environnemental. Un sous-dimensionnement des installations solaires peut entraîner des pertes de production d'énergie, tandis qu'un surdimensionnement représente un investissement initial plus élevé. Par ailleurs, une mauvaise évaluation du potentiel solaire peut conduire à des choix d'implantation non optimaux des centrales solaires, avec des impacts potentiels sur l'environnement.

Pour relever ces défis, de nombreuses méthodes d'estimation du rayonnement solaire ont été développées. Les méthodes empiriques, basées sur des corrélations statistiques entre le rayonnement solaire et des paramètres météorologiques facilement mesurables, sont simples à mettre en œuvre mais présentent des limites en termes de précision. Les modèles physiques, qui simulent les processus de transfert radiatif dans l'atmosphère, offrent une meilleure compréhension des phénomènes physiques mais nécessitent des données d'entrée détaillées souvent difficiles à obtenir. Enfin, les modèles d'apprentissage automatique, notamment les réseaux de neurones artificiels, ont montré un potentiel intéressant pour traiter de grandes quantités de données et établir des relations complexes entre les variables.

1. Le rayonnement solaire : une ressource énergétique majeure

Le rayonnement solaire représente l'une des ressources énergétiques les plus importantes et les plus abondantes sur Terre[1-2]. Il joue un rôle fondamental dans la régulation du climat, les cycles naturels et le fonctionnement des écosystèmes, tout en constituant une

source directe d'énergie renouvelable pour la production d'électricité. Comprendre les composantes et les facteurs qui influencent le rayonnement solaire est essentiel pour optimiser son exploitation [3-5].

1.1.Définition et composantes du rayonnement solaire

Le rayonnement solaire qui atteint la surface de la Terre se compose principalement de trois éléments :

1.1.1. Le rayonnement direct : C'est la portion du rayonnement solaire qui atteint directement la surface terrestre sans être déviée ou absorbée par l'atmosphère. Il est généralement le plus intense sous un ciel dégagé et est essentiel pour les technologies utilisant la concentration solaire, telles que les centrales thermiques à concentration. L'intensité du rayonnement direct est fortement influencée par l'angle d'incidence des rayons solaires et l'épaisseur de l'atmosphère traversée [6-7].

1.1.2. Le rayonnement diffus : C'est la part du rayonnement solaire qui est dispersée par les particules atmosphériques et les molécules d'air. Même sous un ciel partiellement couvert, une partie de ce rayonnement parvient à atteindre la surface terrestre. Le rayonnement diffus est particulièrement important pour les systèmes solaires photovoltaïques à faible concentration ou à suivi solaire limité [8].

1.1.3. Rayonnement réfléchi : Une partie du rayonnement solaire est réfléchi par le sol, la neige, les bâtiments, l'eau, etc. Cette composante dépend fortement de la nature de la surface (albédo). Par exemple, la neige fraîche peut réfléchir jusqu'à 90 % du rayonnement incident. Les albédo-mètres permettent de mesurer cette quantité d'énergie réfléchi.

Ces composantes sont importantes à considérer dans l'évaluation du potentiel solaire d'une région, en particulier dans les zones désertiques comme le Sahara, où les conditions climatiques peuvent significativement affecter leur répartition.

La mesure précise de ces différentes composantes du rayonnement solaire est essentielle pour dimensionner et optimiser les systèmes solaires thermiques et photovoltaïques. Des instruments spécifiques tels que les pyrhéliomètres et les pyranomètres sont utilisés pour mesurer respectivement le rayonnement direct et le rayonnement global [9].

1.2. Le rayonnement global : Il représente la somme totale de l'énergie solaire reçue par une surface horizontale. Il inclut à la fois le rayonnement direct provenant du soleil et le

rayonnement diffus provenant de l'atmosphère. Cette composante globale est souvent utilisée pour évaluer le potentiel solaire d'une région [1].

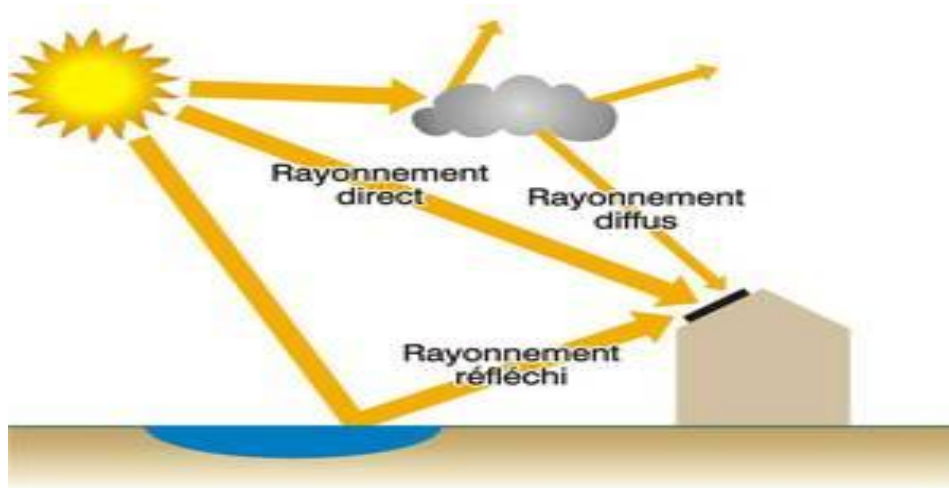


Figure1: Les différents composants du rayonnement solaire.

2. Facteurs influençant le rayonnement solaire

Le rayonnement solaire qui atteint la surface terrestre est influencé par plusieurs facteurs naturels, dont certains varient de manière quotidienne et saisonnière :

2.1.Latitude : Plus une région se situe près de l'équateur, plus l'angle d'incidence des rayons du soleil est direct, ce qui augmente la quantité de rayonnement reçu. À l'inverse, aux latitudes élevées, le rayonnement solaire est plus oblique, diminuant ainsi l'énergie reçue. Cet effet est lié à l'inclinaison de l'axe de rotation de la Terre [1].

2.2.Longitude : est une valeur angulaire qui indique la position d'un point sur Terre, soit à l'est, soit à l'ouest. En géodésie, elle représente l'angle formé entre le plan passant par ce point et l'axe de rotation de la Terre, par rapport au plan du méridien de Greenwich.

Tous les points situés à la même longitude se trouvent sur un demi-plan limité par l'axe des pôles géographiques, découpant la Terre en un demi-cercle centré sur le centre de la planète, reliant ainsi les deux pôles. Ce demi-cercle est désigné sous le nom de méridien.

À l'inverse de la latitude, qui est définie par l'équateur et les pôles, la longitude n'a pas de référence naturelle. Elle est donc exprimée sur une échelle de 360° , s'étendant de -180° à

+180°, soit 180° à l'ouest et 180° à l'est. Le méridien de Greenwich est communément utilisé comme référence pour mesurer la longitude et sert également de base pour la définition des fuseaux horaires.

2.3.Saison : La position de la Terre par rapport au soleil varie au cours de l'année, ce qui modifie l'intensité et la durée de l'ensoleillement. Les régions sahariennes bénéficient d'un rayonnement élevé tout au long de l'année, mais la répartition saisonnière peut également influencer les prévisions et les modèles. L'élongation de la Terre par rapport au Soleil (excentricité orbitale) et l'inclinaison de l'axe de rotation terrestre (obliquité) sont des facteurs supplémentaires à considérer [10].

2.4.Altitude : À mesure que l'altitude augmente, la densité de l'atmosphère diminue, réduisant la quantité d'air et de particules que les rayons solaires doivent traverser. Ainsi, les zones situées en altitude, comme certains plateaux sahariens, peuvent recevoir un rayonnement plus intense. Cependant, l'augmentation de l'albédo (pouvoir réfléchissant) des surfaces à haute altitude peut compenser en partie cet effet [11].

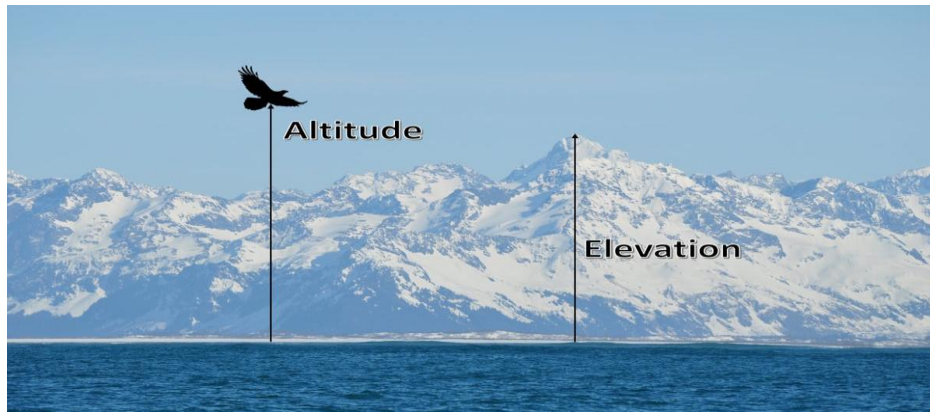


Figure 2: l'altitude.

2.5.Couverture nuageuse

Les nuages agissent comme des écrans partiels ou totaux du rayonnement solaire. Dans les régions désertiques où la couverture nuageuse est généralement faible, le rayonnement direct est souvent dominant, mais des phénomènes occasionnels comme les tempêtes de sable peuvent temporairement diminuer l'intensité du rayonnement. La nature des nuages (cirrus, cumulus, stratus) influence différemment la transmission du rayonnement [12].

2.6.Transparence atmosphérique et aérosols

La transparence de l'atmosphère, c'est-à-dire sa capacité à laisser passer les rayons du soleil, est influencée par la présence de particules, de gaz et de poussières, appelés aérosols. Dans les zones désertiques comme le Sahara, les tempêtes de sable et la présence d'aérosols en suspension peuvent réduire la quantité de rayonnement direct et augmenter le rayonnement diffus. La composition chimique des aérosols et leur taille influencent leur capacité d'absorption et de diffusion du rayonnement [13].

2.7.Activité solaire

Les variations de l'activité solaire, mesurées par le nombre de taches solaires, influencent légèrement le rayonnement solaire atteignant la Terre. Cependant, cet effet est généralement considéré comme négligeable par rapport aux autres facteurs [14].

3. Le rayonnement solaire : un moteur de la vie sur Terre

Le rayonnement solaire est un élément clé dans les bilans énergétiques terrestres et joue un rôle fondamental dans de nombreux processus naturels :

3.1.Cycle de l'eau

Le rayonnement solaire est à l'origine de l'évaporation de l'eau des océans, des lacs et des rivières, initiant ainsi le cycle de l'eau. Ce processus d'évaporation est essentiel pour la formation des nuages et la distribution des précipitations à l'échelle mondiale. L'énergie solaire fournit l'énergie latente nécessaire à la transformation de l'eau liquide en vapeur [15].

3.2.Climat

Le rayonnement solaire influe directement sur la température de l'atmosphère et de la surface terrestre, ce qui régule le climat global. Il est également un moteur des vents et des courants océaniques, influençant ainsi les conditions météorologiques locales et mondiales. Les variations de l'insolation solaire au cours de l'année et les interactions complexes entre le rayonnement solaire et les gaz à effet de serre jouent un rôle crucial dans les changements climatiques [16].

3.3.Écosystèmes

La photosynthèse, processus vital pour la production de biomasse, dépend entièrement du rayonnement solaire. C'est grâce à cette énergie que les plantes peuvent transformer le

dioxyde de carbone et l'eau en oxygène et en matière organique, assurant ainsi la base de la chaîne alimentaire. Le rayonnement solaire est donc à l'origine de la diversité biologique sur Terre [17].

En outre, le rayonnement solaire joue un rôle central dans le bilan énergétique des bâtiments et des systèmes de production d'énergie. Maximiser son utilisation est un objectif fondamental dans la transition vers des systèmes énergétiques plus durables et respectueux de l'environnement. Les technologies solaires photovoltaïques et thermiques permettent de convertir directement le rayonnement solaire en électricité ou en chaleur, contribuant ainsi à réduire notre dépendance aux énergies fossiles [18].

4. Les méthodes traditionnelles d'estimation du rayonnement solaire

L'estimation du rayonnement solaire est un défi fondamental dans de nombreux domaines, allant de l'énergie solaire à la météorologie en passant par la gestion des ressources environnementales. Plusieurs méthodes ont été développées au fil du temps pour évaluer cette ressource, et elles peuvent être classées en trois grandes catégories : les modèles empiriques, les modèles physiques, et les modèles basés sur l'intelligence artificielle. Chacune de ces approches présente des avantages et des limites en fonction des besoins et des conditions locales.

4.1.Modèles empiriques

Les chercheurs dans le domaine solaires ont développé de nombreuses corrélations empiriques qui déterminent la relation entre le rayonnement solaire et autres paramètres météorologiques basant sur la disponibilité des informations météorologiques. Ces paramètres sont utilisés comme des entrées dans les modèles de rayonnement. Les modèles empiriques peuvent être principalement classés en quatre catégories basées sur les paramètres météorologiques [19].

- Modèles basés sur l'ensoleillement.
- Modèles basés sur le Cloud.
- Modèles basés sur la température.
- Autres modèles basés sur des paramètres météorologiques.

❖ **Avantages [19] :**

- **Simplicité :** Ces modèles sont faciles à mettre en œuvre et ne nécessitent généralement que quelques paramètres d'entrée, tels que les données de température ou de précipitation.
- **Peu de données requises :** Contrairement aux modèles plus complexes, les modèles empiriques peuvent fonctionner avec un ensemble minimal de données, ce qui est un avantage majeur dans les régions où les mesures météorologiques sont limitées.

❖ **Limites [19]:**

- **Validité locale :** Les modèles empiriques sont souvent calibrés pour une région ou un climat particulier, ce qui limite leur capacité à être généralisés à d'autres environnements. Leur performance peut se dégrader lorsque les conditions climatiques diffèrent de celles pour lesquelles ils ont été développés.
- **Sensibilité aux paramètres :** Ces modèles dépendent fortement des paramètres choisis et peuvent produire des résultats imprécis si les valeurs de ces paramètres ne sont pas adéquates.
- **Incapacité à représenter des phénomènes complexes :** Les modèles empiriques ne prennent généralement pas en compte les interactions complexes entre l'atmosphère et le rayonnement solaire, ce qui limite leur précision.

• **Exemples :**

Les modèles basés sur l'ensoleillement sont parmi les plus utilisés pour estimer le rayonnement solaire en raison de la disponibilité des données et de leur simplicité. Ces modèles établissent une relation entre le rayonnement solaire global et la durée d'insolation, souvent exprimée sous forme de fraction par rapport à la durée maximale possible (indice de clarté).

Formules et Approches Courantes**1. Modèle d'Angström (1924) – Modifié par Prescott (1940)[20-21]**

$$\frac{H}{H_0} = a + b \left(\frac{S}{S_0} \right) \quad (1)$$

- H : Rayonnement solaire global reçu.
- H_0 : Rayonnement extraterrestre.

- S : Durée réelle d'ensoleillement.
- S_0 : Durée maximale possible d'ensoleillement.
- a, b : Coefficients empiriques dépendant du climat local.

2. Modèle de Hargreaves & Samani (1982) [22-23]

$$H = K \cdot (T_{max} - T_{min})^{0.5} \cdot H_0 \quad (2)$$

- Intègre l'amplitude thermique comme proxy de l'ensoleillement.
- K : Coefficient ajustable ($\approx 0,16$ pour les régions intérieures).

3. Modèles Régionaux (Adaptations Locales)

Des variantes incluent des corrections pour l'humidité, l'altitude, ou la nébulosité moyenne.

4.2. Modèles Physiques de Rayonnement Solaire : Principes et Mécanismes

Les modèles physiques (ou déterministes) estiment le rayonnement solaire en simulant les processus atmosphériques qui affectent la propagation de la lumière du soleil. Contrairement aux modèles empiriques (basés sur des corrélations statistiques), ils reposent sur des équations fondamentales de la physique radiative, offrant une approche plus universelle mais plus complexe [24].

Processus Clés Simulés

1. Atténuation par les gaz atmosphériques [25]

- Absorption : Certaines longueurs d'onde sont absorbées par des gaz comme l'ozone (O_3), le dioxyde de carbone (CO_2) ou la vapeur d'eau (H_2O).
- Loi de Beer-Lambert : Décrit l'affaiblissement du rayonnement en fonction de la concentration des gaz et de la distance parcourue.

2. Diffusion par les aérosols et les molécules

- Diffusion de Rayleigh (molécules d'air) : Domine par temps clair (ciel bleu).
- Diffusion de Mie (aérosols, poussières) : Influence la nébulosité et la pollution.

3. Interaction avec les nuages

- Réflexion, absorption et transmission variant selon :
 - Épaisseur optique des nuages.
 - Taille des gouttelettes/cristaux.
 - Altitude (nuages bas vs hauts).

4. Effets de surface

- Albédo : Fraction de rayonnement réfléchi par le sol (neige, désert, forêt...).
- Angle d'incidence : Dépend de la latitude, de l'heure et de la saison.

Exemples de Modèles Physiques

- **MODTRAN (MODerate resolution TRANsmission)** : Utilisé pour les applications militaires et satellitaires [26].
- **libRadtran** : Bibliothèque open-source pour le transfert radiatif [27].
- **RRTM (Rapid Radiative Transfer Model)** : Intégré dans les modèles climatiques comme le CESM [28].

Avantages

- **Précision élevée** : Capture les variations fines (ex. impact des aérosols volcaniques).
- **Universalité** : Fonctionne sous tous les climats sans recalibration locale.
- **Capacité prédictive** : Couplage possible avec les modèles météo/climat.

Limitations

- **Complexité** : Requiert des données détaillées (ex. profils verticaux d'humidité).
- **Coût calcul** : Simulations gourmandes en ressources (supercalculateurs souvent nécessaires).
- **Sensibilité aux inputs** : Erreurs dans les données atmosphériques dégradent les résultats.

Applications

- **Énergie solaire** : Optimisation des centrales photovoltaïques/concentrées.

- Télédétection : Correction des images satellitaires.
- Climatologie : Étude du forçage radiatif (ex. rôle des nuages dans le réchauffement).

4.3. Modèles basés sur l'intelligence artificielle

Les modèles basés sur l'intelligence artificielle (IA) révolutionnent la prédiction du rayonnement solaire grâce à leur capacité à apprendre des relations complexes à partir de données. Contrairement aux modèles empiriques ou physiques, ils ne reposent pas sur des équations prédéfinies, mais sur des algorithmes capables de s'adapter et d'améliorer leurs performances avec l'expérience.

1. Principes des Modèles IA pour le Rayonnement Solaire

1.1. Architectures Couramment Utilisées

◇ Réseaux de Neurones Multicouches (MLP - *Multi layer Perceptron*) [29]

Fonctionnement :

- Structure feedforward avec couches d'entrée, cachées et de sortie.
- Chaque neurone applique une fonction d'activation (*ReLU*, *sigmoïde*) pour modéliser des relations non linéaires.

Équation d'un Neurone :

$$y = f(\sum_{i=1}^n (w_i x_i) + b) \quad (3)$$

Avec : w_i : Poids, x_i : Entrées, b : Biais, f : Fonction d'activation.

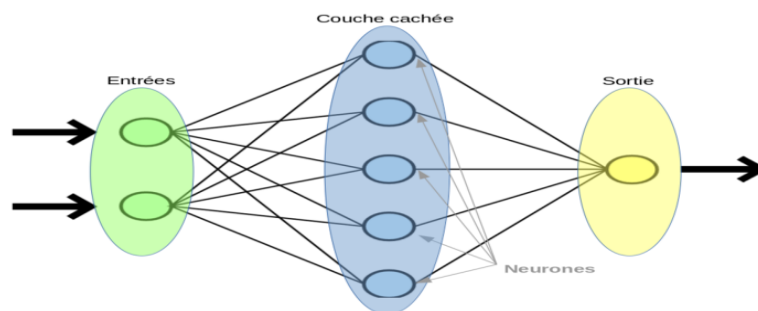


Figure 3 : Structure de base d'un MLP

Applications :

- Prédiction horaire/journalière du rayonnement.
- Combinaison de données météo (température, humidité, nébulosité).

◇ Réseaux de Neurones Convolutifs (CNN - *Convolutional Neural Networks*) [30]

Fonctionnement :

- Conçus pour traiter des données spatiales (ex. images satellitaires de couverture nuageuse).
- Utilisent des kernels pour extraire des motifs locaux.

Équation de Convolution :

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i - m, j - n) \cdot K(m, n) \quad (4)$$

- I : Image d'entrée, K : Noyau de convolution.

Applications :

- Estimation du rayonnement à partir de cartes de nuages.
- Prévision à très haute résolution (énergie solaire distribuée).

◇ ANFIS (*Adaptive Neuro-Fuzzy Inference System*) [31-32]

Fonctionnement :

- Combine réseaux de neurones et logique floue pour modéliser l'incertitude.
- 5 couches : Fuzzification, règles floues, normalisation, inférence, défuzzification.

Équation (Règle Floue Type Takagi-Sugeno) :

$$y = \frac{\sum_{i=1}^n w_i \cdot f_i(x)}{\sum_{i=1}^n w_i} \quad (5)$$

- w_i : Poids de la règle, f_i : Fonction linéaire.

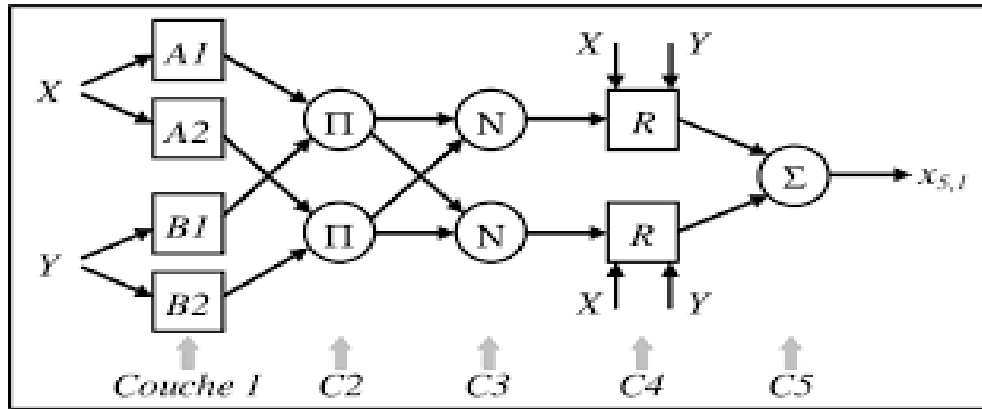


Figure 4: Illustration des couches ANFIS avec entrées floues et sortie précise.

Applications :

- Modélisation dans des environnements variables (ex. zones tropicales humides).

Avantages et Limites des Modèles IA

Avantages

- **Adaptabilité** : Capture des relations complexes (non-linéarités, interactions entre variables).
- **Performance** : Surpasse souvent les modèles empiriques avec suffisamment de données.
- **Automatisation** : Peut s'améliorer continuellement avec de nouvelles données (*online learning*).

Limites

- **Besoins en données** : Requiert des data sets massifs et bien annotés.
- **Opacité (Black Box)** : Difficile d'interpréter comment la prédiction est générée.
- **Coût calcul** : Entraînement gourmand en ressources (GPU nécessaires pour les CNN/MLP profonds).

Tableau 1: Comparaison des Modèles IA vs. Empiriques/Physiques

<i>Critère</i>	<i>Modèles IA</i>	<i>Modèles Empiriques</i>	<i>Modèles Physiques</i>
<i>Précision</i>	<i>Très levées (si données suffisantes)</i>	<i>Moyenne (dépend des paramètres)</i>	<i>Élevée (si données exactes)</i>
<i>Besoins en données</i>	<i>Volume élevé</i>	<i>Faible (ensoleillement)</i>	<i>Données atmosphériques détaillées</i>
<i>Interprétabilité</i>	<i>Faible (boîte noire)</i>	<i>Moyenne (équations simples)</i>	<i>Élevée (lois physiques)</i>
<i>Temps de calcul</i>	<i>Lent (entraînement)</i>	<i>Rapide</i>	<i>Très lent (supercalculateurs)</i>

5. Spécificités du climat saharien et impact sur l'estimation du rayonnement solaire

Le Sahara est l'une des régions les plus arides et les plus ensoleillées au monde. Cependant, ses conditions climatiques extrêmes posent des défis particuliers pour l'estimation précise du rayonnement solaire. L'analyse des spécificités du climat saharien permet de mieux comprendre comment ces caractéristiques influencent l'estimation du rayonnement solaire, et de proposer des solutions pour surmonter ces obstacles.

5.1. Caractéristiques climatiques du Sahara [33]

Le climat saharien présente des particularités uniques qui influencent directement le rayonnement solaire reçu à la surface :

- **Fort ensoleillement annuel** : Le Sahara bénéficie d'un des taux d'ensoleillement les plus élevés au monde, avec une moyenne de 4 000 à 4 500 heures d'ensoleillement par an. Cela en fait une région idéale pour exploiter l'énergie solaire.
- **Amplitude thermique élevée** : Les températures dans le Sahara peuvent varier considérablement au cours de la journée, atteignant des extrêmes de plus de 50 °C pendant la journée et descendant sous 0 °C la nuit. Cette forte amplitude thermique peut

influencer les conditions atmosphériques, telles que la turbulence, qui affectent à leur tour le rayonnement solaire.

- **Faible humidité** : L'air dans le Sahara est extrêmement sec, avec une humidité relative très faible, ce qui permet une meilleure transmission du rayonnement solaire direct. Toutefois, cette sécheresse entraîne également des phénomènes de poussière fréquents.
- **Phénomènes de poussières et de sable** : Les tempêtes de sable et de poussière sont fréquentes dans le désert, surtout pendant certaines saisons. Ces particules en suspension dans l'atmosphère peuvent significativement atténuer le rayonnement solaire et affecter sa répartition entre les composantes directe et diffuse.
- **Vent fort et régulier** : Les vents dans le Sahara sont souvent forts et réguliers, notamment les vents d'Est comme le sirocco. Ces vents soulèvent des particules de sable, aggravant les phénomènes de poussière et modifiant les conditions atmosphériques, influençant ainsi le transfert radiatif.

5.2. Conséquences sur le rayonnement solaire

Les spécificités du climat saharien ont des effets marqués sur le rayonnement solaire, notamment sur sa répartition et son intensité :

- **Variation saisonnière importante** : Bien que le Sahara bénéficie d'un ensoleillement presque constant tout au long de l'année, des variations saisonnières existent en raison des changements d'angle d'incidence des rayons du soleil et de la couverture nuageuse limitée. Ces variations peuvent affecter la quantité de rayonnement solaire disponible à différentes périodes de l'année.
- **Effet des poussières sur l'atténuation du rayonnement** : Les tempêtes de sable et les poussières en suspension dans l'atmosphère saharienne contribuent à une forte atténuation du rayonnement solaire direct, augmentant la proportion de rayonnement diffus. Ces particules dispersent et absorbent une partie du rayonnement, ce qui peut réduire l'intensité du rayonnement global qui atteint le sol.
- **Influence de la turbulence atmosphérique** : La forte amplitude thermique entraîne une turbulence atmosphérique significative, notamment à la surface du sol. Cette turbulence modifie les conditions de transfert radiatif et peut générer des fluctuations dans le rayonnement solaire reçu.

6. Difficultés liées à l'estimation du rayonnement solaire dans le Sahara

Les caractéristiques climatiques du Sahara posent plusieurs défis pour l'estimation précise du rayonnement solaire dans cette région :

- **Manque de stations météorologiques :** Le Sahara, en raison de son étendue et de ses conditions difficiles, est faiblement instrumenté. Le nombre limité de stations météorologiques rend difficile la collecte de données complètes et précises sur le rayonnement solaire et d'autres variables atmosphériques, ce qui complique le calibrage des modèles.
- **Qualité des données souvent limitée :** Les données météorologiques disponibles dans le Sahara peuvent être incomplètes, intermittentes, ou de qualité inférieure en raison des conditions extrêmes et de l'éloignement des sites de mesure. Ces données sont souvent insuffisantes pour alimenter des modèles de transfert radiatif ou des méthodes d'intelligence artificielle avec précision.
- **Variabilité spatio-temporelle élevée :** Le Sahara est une région hétérogène où les conditions climatiques peuvent varier rapidement à la fois dans l'espace et dans le temps. Cette variabilité rend l'estimation du rayonnement solaire plus difficile car les modèles doivent tenir compte des changements fréquents dans les conditions atmosphériques, tels que les tempêtes de sable, les variations de température et les fluctuations de l'humidité.

Les conditions climatiques extrêmes du Sahara, bien que favorables à une exploitation massive de l'énergie solaire, posent des défis spécifiques pour l'estimation précise du rayonnement solaire. Les phénomènes de poussières, l'amplitude thermique et la variabilité spatiale des conditions météorologiques exigent des méthodes de modélisation robustes et adaptées à ce contexte unique. Le développement de modèles prenant en compte ces spécificités est essentiel pour maximiser le potentiel solaire du Sahara tout en minimisant les incertitudes dans les prévisions.

7. Les limites des modèles existants dans le contexte saharien

Le Sahara, avec ses conditions climatiques extrêmes, pose des défis uniques pour l'estimation précise du rayonnement solaire. Bien que de nombreux modèles aient été développés pour estimer cette ressource, leur application dans un environnement aussi particulier révèle des limites significatives. Cette section examine les principales limites des modèles empiriques, physiques et basés sur l'intelligence artificielle dans le contexte saharien.

Modèles empiriques : une inadéquation avec les conditions extrêmes du Sahara

Les modèles empiriques, largement utilisés pour l'estimation du rayonnement solaire, sont généralement calibrés à partir de données locales et sont efficaces dans des contextes

climatiques stables. Cependant, ces modèles présentent plusieurs limites lorsqu'ils sont appliqués au climat saharien.

- **Peu adaptés aux conditions extrêmes :** Les modèles empiriques, tels que les modèles d'Angström ou de Hargreaves-Samani, sont souvent conçus pour des conditions climatiques modérées. Le Sahara, avec ses températures extrêmes, ses phénomènes de poussières fréquents et son ensoleillement quasi constant, dépasse les conditions pour lesquelles ces modèles sont optimisés. Cela peut conduire à des erreurs importantes dans les estimations de rayonnement solaire.
- **Simplification excessive des phénomènes atmosphériques :** Les modèles empiriques reposent sur des relations statistiques simples entre quelques variables, comme la température et l'ensoleillement. Or, dans le Sahara, des facteurs complexes tels que la concentration en aérosols, la turbulence atmosphérique et les tempêtes de sable jouent un rôle majeur dans l'atténuation du rayonnement solaire, des éléments que les modèles empiriques ne capturent pas adéquatement.
- **Validité locale limitée :** Comme ces modèles sont généralement calibrés pour des conditions locales spécifiques, leur application au Sahara, qui possède des caractéristiques climatiques très différentes de celles des zones où ils ont été développés, peut entraîner des prévisions biaisées et moins fiables.

Modèles physiques : une exigence de données inaccessibles

Les modèles physiques de transfert radiatif sont, en théorie, parmi les plus précis pour estimer le rayonnement solaire. Cependant, leur utilisation dans le Sahara est entravée par plusieurs obstacles liés aux spécificités des données requises.

- **Nécessité de données d'entrée spécifiques :** Les modèles de transfert radiatif, tels que MODTRAN ou DISORT, nécessitent des données d'entrée très détaillées, notamment sur la composition atmosphérique, les profils d'aérosols, l'humidité, et la couverture nuageuse. Dans le désert, ces données sont souvent rares ou inexistantes, en raison du faible nombre de stations météorologiques et de l'éloignement géographique des infrastructures de mesure.
- **Inaccessibilité des mesures atmosphériques précises :** Les phénomènes tels que les tempêtes de sable et les poussières en suspension, qui influencent fortement le rayonnement solaire dans le Sahara, nécessitent des données sur les concentrations d'aérosols et leur composition. Or, ces informations ne sont pas systématiquement

mesurées dans cette région, ce qui limite la capacité des modèles physiques à capturer ces effets avec précision.

- **Ressources computationnelles importantes :** Les modèles physiques nécessitent des temps de calcul et des ressources de traitement considérables, ce qui peut être un frein dans des contextes où l'accès aux infrastructures de calcul est limité. Cela rend leur application en temps réel ou à grande échelle difficile à envisager dans un environnement comme le Sahara.

Modèles d'intelligence artificielle : un potentiel limité par le manque de données

Les modèles basés sur l'intelligence artificielle (IA) ont montré des résultats prometteurs pour l'estimation du rayonnement solaire, notamment grâce à leur capacité d'apprentissage et à leur flexibilité. Cependant, dans le contexte saharien, leur efficacité est souvent restreinte par le manque de données d'apprentissage de qualité.

- **Dépendance aux données d'apprentissage :** Les modèles d'IA, tels que les réseaux de neurones ou les machines à vecteurs de support, nécessitent de grands ensembles de données d'apprentissage pour être entraînés correctement. Dans le Sahara, le manque de stations météorologiques et la rareté des données historiques sur le rayonnement solaire limitent la quantité de données disponibles pour l'entraînement de ces modèles, réduisant ainsi leur précision.
- **Données souvent de faible qualité :** En plus de la rareté des données, celles qui sont disponibles sont parfois incomplètes ou bruitées, ce qui peut nuire aux performances des modèles d'IA. Les modèles risquent d'apprendre des relations incorrectes ou peu généralisables si les données sont de mauvaise qualité ou si elles ne reflètent pas correctement la variabilité spatio-temporelle du climat saharien.
- **Risque de surapprentissage :** Dans un contexte où les données sont limitées, les modèles d'IA peuvent souffrir de surapprentissage, c'est-à-dire qu'ils s'adaptent trop spécifiquement aux données d'entraînement, au détriment de leur capacité à généraliser à d'autres périodes ou régions. Ce problème est particulièrement pertinent dans le Sahara, où les conditions climatiques sont souvent imprévisibles et sujettes à de fortes variations.

8. Les méta-heuristiques : une approche prometteuse

Dans le domaine de l'estimation du rayonnement solaire, les méthodes traditionnelles, qu'elles soient empiriques, physiques ou basées sur l'intelligence artificielle, présentent des limitations dans des environnements complexes comme celui du Sahara. Les méta-heuristiques, inspirées par des processus naturels et évolutifs, offrent une alternative innovante et puissante pour surmonter ces défis. Cette section explore les principes des méta-heuristiques et leurs avantages pour l'estimation du rayonnement solaire dans un contexte saharien.

• Définition et principes des méta-heuristiques

Les méta-heuristiques sont des algorithmes d'optimisation puissants, inspirés par des processus observés dans la nature, qui cherchent à résoudre des problèmes complexes où les approches traditionnelles échouent ou sont inefficaces. Ces algorithmes permettent d'explorer efficacement l'espace de recherche pour identifier des solutions optimales ou quasi-optimales dans des problèmes multidimensionnels et non linéaires.

Algorithmes d'optimisation inspirés de la nature : Les méta-heuristiques incluent plusieurs familles d'algorithmes inspirés de phénomènes naturels, tels que [34] :

Les algorithmes génétiques (GA), basés sur la sélection naturelle et l'évolution biologique, qui explorent les solutions par mutation et recombinaison.

Les algorithmes de recherche tabou, qui interdisent temporairement certaines solutions pour éviter les boucles locales et améliorer l'exploration de l'espace de recherche.

Les essaims de particules (PSO), qui simulent le comportement des essaims ou des troupes d'animaux, chaque particule représentant une solution potentielle qui s'adapte en fonction de son propre parcours et de celui des autres particules.

Les algorithmes de colonies de fourmis, qui s'inspirent du comportement coopératif des fourmis pour trouver des chemins optimaux entre plusieurs points.

9. Exploration efficace de l'espace de recherche

Contrairement aux méthodes classiques d'optimisation qui peuvent se bloquer dans des optima locaux, les méta-heuristiques sont conçues pour éviter ce type de problème en explorant différentes régions de l'espace de recherche. Elles peuvent naviguer efficacement

dans des espaces complexes, même lorsque les relations entre les variables sont non linéaires ou mal définies.

9.1. Avantages des méta-heuristiques pour l'estimation du rayonnement solaire

L'estimation du rayonnement solaire, en particulier dans des environnements difficiles comme le Sahara, pose de nombreux défis en raison de la complexité des phénomènes atmosphériques, de la variabilité des conditions climatiques, et du manque de données. Les méta-heuristiques présentent plusieurs avantages clés qui les rendent particulièrement adaptées à cette tâche :

Capacité à gérer des problèmes complexes et non linéaires : Le rayonnement solaire est influencé par une multitude de facteurs interdépendants, tels que les aérosols, la couverture nuageuse, la température et la latitude. Les méta-heuristiques sont capables de modéliser des relations non linéaires complexes et de traiter des ensembles de données hétérogènes. Cela leur permet de fournir des estimations plus robustes et précises dans des contextes où les modèles traditionnels échouent.

Robustesse face au manque de données : Dans des régions comme le Sahara, où les stations météorologiques sont peu nombreuses et où les données sont souvent fragmentaires ou bruitées, les méta-heuristiques peuvent exploiter les informations disponibles de manière efficace. Grâce à leur flexibilité, elles permettent de compenser partiellement le manque de données en adaptant leur exploration de l'espace de recherche en fonction des informations disponibles, tout en évitant les surapprentissage liés à des jeux de données incomplets.

Flexibilité pour combiner différentes méthodes : L'un des principaux avantages des méta-heuristiques est leur capacité à être intégrées à d'autres méthodes d'estimation, telles que les réseaux de neurones, les modèles empiriques, ou les modèles physiques. En les combinant, on peut améliorer les performances globales de l'estimation du rayonnement solaire, en exploitant les points forts de chaque méthode pour pallier leurs limites respectives. Par exemple, un algorithme d'optimisation par essaims de particules peut être utilisé pour ajuster les paramètres d'un modèle de réseau de neurones, améliorant ainsi la précision des prédictions dans un environnement complexe.

9.2. Exemples d'applications en météorologie

Les méta-heuristiques ont déjà montré leur efficacité dans divers domaines de la météorologie, notamment pour optimiser les modèles numériques de prévision et assimiler

les données atmosphériques dans des conditions variables. Voici quelques exemples d'applications :

Optimisation des paramètres de modèles numériques : Les méta-heuristiques sont utilisées pour ajuster les paramètres des modèles de prévision numérique du temps, en tenant compte des incertitudes liées aux conditions initiales et aux variations atmosphériques locales. Cela permet de maximiser la précision des prédictions dans des environnements difficiles comme les zones désertiques [35].

Assimilation de données : Dans des régions où les données météorologiques sont rares ou imprécises, les algorithmes méta-heuristiques peuvent être utilisés pour assimiler des observations provenant de sources variées (stations terrestres, satellites, etc.) et combler les lacunes dans les modèles de prévision. Cela est particulièrement pertinent pour l'estimation du rayonnement solaire dans des régions mal couvertes comme le Sahara [36].

Prédiction des événements météorologiques extrêmes : Les méta-heuristiques ont été appliquées pour prévoir des événements météorologiques extrêmes, tels que les tempêtes de sable, en optimisant les modèles physiques en fonction des données disponibles. Leur capacité à gérer des systèmes dynamiques et complexes est un atout majeur pour la météorologie désertique [37].

Conclusion

En résumé, ce chapitre a souligné l'importance stratégique de l'énergie solaire, en particulier dans le contexte saharien, où les conditions climatiques offrent un potentiel énergétique exceptionnel. Cependant, il a également mis en évidence les défis spécifiques à l'estimation du rayonnement solaire dans cette région, tels que la variabilité spatio-temporelle, le manque de données de qualité, et les phénomènes atmosphériques complexes comme les tempêtes de sable. Les méthodes traditionnelles d'estimation, qu'elles soient empiriques, physiques ou basées sur l'intelligence artificielle, présentent des limites importantes face à ces défis.

L'exploration des méta-heuristiques a montré leur capacité à pallier certaines de ces insuffisances grâce à leur flexibilité, leur robustesse face au manque de données et leur aptitude à modéliser des problèmes complexes et non linéaires. Cependant, ces méthodes ne sont pas exemptes de limitations, ce qui justifie la nécessité d'une approche hybride combinant

différentes techniques pour tirer parti de leurs avantages respectifs tout en minimisant leurs faiblesses.

Dans cette optique, le développement d'un modèle hybride apparaît comme une solution prometteuse pour améliorer l'estimation du rayonnement solaire dans le Sahara. Ce modèle, en combinant les méta-heuristiques avec des méthodes traditionnelles ou des techniques d'intelligence artificielle, pourrait offrir une précision accrue et une meilleure capacité à gérer la complexité du climat saharien.

Le chapitre suivant sera consacré à la présentation de cette approche hybride innovante. Il détaillera la structure du modèle proposé, les techniques employées, ainsi que les résultats attendus pour l'estimation du rayonnement solaire dans les conditions extrêmes du Sahara.

Références

- [1] Iqbal, M. (1983). An introduction to solar radiation. Academic Press.
- [2] Benkouider, F., & Gueymard, C. (2005). Estimation of the global solar radiation over Algeria using satellite images. *Solar Energy*, 79(1), 1-10.
- [3] Benatallah, D., Bouchouicha, K., Benatallah, A., Harrouz, A., & Nasri, B. (2019). Forecasting of solar radiation using an empirical model. *Algerian Journal of Renewable Energy and Sustainable Development*, 1(2), 212-219.
- [4] Angström, A. (1924). A simple method of computing the global radiation of the sun. *Geographical Review*, 14(2), 313-321.
- [5] Gueymard, C. (2005). SMARTS2: A simple model for the atmospheric radiative transfer of sunshine. *Solar Energy*, 79(1), 11-23.
- [6] Kambezidis, H. D., & Asimakopoulou, D. N. (2004). Artificial neural networks for the prediction of global solar radiation. *Solar Energy*, 77(6), 457-467.
- [7] Riou, C., Valancogne, C., & Pieri, P. (1989). Un modèle simple d'interception du rayonnement solaire par la vigne-vérification expérimentale. *Agronomie*, 9(5), 441-450.
- [8] Duffie, J. A., & Beckman, W. A. (1980). *Solar engineering of thermal processes* (p. 16591). New York: Wiley.
- [9] Martin, M., & Bourges, B. (2002). Instrumentation pour la mesure du rayonnement solaire. *Techniques de l'Ingénieur*.
- [10] North, G. R., & Tang, W. (1989). Variations in solar radiation and climate. *Science*, 246(4926), 499-503.
- [11] Pinker, R. T., & Laszlo, I. (1992). Modeling surface solar radiation for satellite remote sensing. *Journal of Geophysical Research: Atmospheres*, 97(D14), 14529-14550.
- [12] Stephens, G. L. (2005). Cloud droplet nucleation: a critical review of observations and theory. *Atmospheric Research*, 76(1-4), 71-98.

-
- [13] Seinfeld, J. H., & Pandis, S. N. (2016). *Atmospheric chemistry and physics: From air pollution to climate change*. John Wiley & Sons.
 - [14] Hoyt, D. V., & Schatten, K. H. (1993). The role of solar variability in climate change. *Reviews of Geophysics*, 31(4), 403-440.
 - [15] Trenberth, K. E. (1976). Spatial and temporal variations of the atmospheric water balance. *Journal of Geophysical Research*, 81(21), 3657-3667.
 - [16] Hansen, J., Sato, M., Ruedy, R., Lacis, A., & Oeschger, H. (1998). Global sea surface temperature change: Is it reversing? *Science*, 281(5379), 1421-1425.
 - [17] Coquillard, P., & Hill, D. R. (1997). *Modélisation et simulation des écosystèmes* (pp. 273-p).
 - [18] Denholm, P., & Margolis, R. M. (2007). *Evolution of solar electric power systems in the United States*. National Renewable Energy Laboratory.
 - [19] Besharat, F., Dehghan, A. A., & Faghih, A. R. (2013). Empirical models for estimating global solar radiation: A review and case study. *Renewable and Sustainable Energy Reviews*, 21, 798-821.
 - [20] Angstrom A. Solar and terrestrial radiation. *Quarterly Journal of Royal Meteorological Society* 1924;50:121–5.
 - [21] Prescott JA. Evaporation from water surface in relation to solar radiation. *Transactions of the Royal Society of Australia* 1940;46:114–8.
 - [22] Hargreaves GH, Samani ZA. Estimating potential evapotranspiration. *Journal of Irrigation and Drainage Engineering* 1982;108(IR3):223–30.
 - [23] Hargreaves GH. Simplified coefficients for estimating monthly solar radiation in North America and Europe. Departmental paper, Departmental of Biological and Irrigation Engineering, Utah State University, Logan; 1994.
 - [24] Diabaté, L. (1989). *Détermination du rayonnement solaire à l'aide d'images satellitaires* (Doctoral dissertation, École Nationale Supérieure des Mines de Paris).
 - [25] Akinwumi, S. A., Omotosho, T. V., Usikal, M. R., Adewusi, M. O., & Ometan, O. O. (2016, October). Atmospheric gases attenuation in West Africa. In *2016 IEEE Radio and Antenna Days of the Indian Ocean (RADIO)* (pp. 1-2). IEEE.
 - [26] Ballestrín, J., & Marzo, A. (2012). Solar radiation attenuation in solar tower plants. *Solar Energy*, 86(1), 388-392.
 - [27] Mayer, B., Kylling, A., Emde, C., Hamann, U., & Buras, R. (2012). *libRadtran user's guide*. Edition for libRadtran version, 1.
 - [28] Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., ... & Brown, P. D. (2005). Atmospheric radiative transfer modeling: A summary of the AER codes. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 91(2), 233-244.
 - [29] Taud, H., & Mas, J. F. (2017). Multilayer perceptron (MLP). In *Geomatic approaches for modeling land change scenarios* (pp. 451-455). Cham: Springer International Publishing.
 - [30] O'shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.
 - [31] Walia, N., Singh, H., & Sharma, A. (2015). ANFIS: Adaptive neuro-fuzzy inference system-a survey. *International Journal of Computer Applications*, 123(13).

- [32] Salleh, M. N. M., Talpur, N., & Hussain, K. (2017). Adaptive neuro-fuzzy inference system: Overview, strengths, limitations, and solutions. In *Data Mining and Big Data: Second International Conference, DMBD 2017, Fukuoka, Japan, July 27–August 1, 2017, Proceedings 2* (pp. 527-535). Springer International Publishing.
- [33] Boumaraf, B. (2015). *Caractéristiques et fonctionnement des sols dans la vallée d'Oued Righ, Sahara Nord oriental, Algérie* (Doctoral dissertation, Université Mohamed Khider-Biskra).
- [34] Boussaid, I. (2013). *Perfectionnement de métaheuristiques pour l'optimisation continue* (Doctoral dissertation, Université Paris-Est; Université des Sciences et de la Technologie Houari-Boumediène (Algérie)).
- [35] Jiang, H., Wang, J., Wu, J., & Geng, W. (2017). Comparison of numerical methods and metaheuristic optimization algorithms for estimating parameters for wind energy potential assessment in low wind regions. *Renewable and Sustainable Energy Reviews*, 69, 1199-1217.
- [36] Cintra, R. S. C., Cocke, S., & Campos Velho, H. F. (2018). Data assimilation by neural networks with ensemble prediction. In *Proceedings of the International Symposium on Uncertainty Modeling and Analysis, Florianópolis (SC), Brazil* (pp. 35-43).
- [37] Jovanovic, L., Bacanin, N., Simic, V., Mani, J., Zivkovic, M., & Sarac, M. (2024). Optimizing machine learning for space weather forecasting and event classification using modified metaheuristics. *Soft Computing*, 28(7), 6383-6402.



Chapitre III :

Méthodologie et outils



Introduction

L'estimation précise du rayonnement solaire est une étape essentielle pour diverses applications en énergétique, notamment dans le dimensionnement des systèmes photovoltaïques et thermiques[1]. Une bonne connaissance du rayonnement solaire permet d'optimiser la production d'énergie et d'améliorer l'efficacité des systèmes exploitant cette ressource naturelle. Cependant, en raison de la variabilité spatiale et temporelle du rayonnement solaire, il est indispensable d'adopter une approche méthodologique rigoureuse qui intègre à la fois des données fiables, des modèles prédictifs performants et des outils de traitement avancés[2].

Ce chapitre présente la méthodologie adoptée pour la collecte et le prétraitement des données météorologiques, ainsi que le développement des modèles d'estimation du rayonnement solaire.

Tout d'abord, une description des différentes sources de données utilisées est effectuée, incluant les données météorologiques terrestres issues des stations de mesure, les données satellitaires fournies par divers organismes internationaux, et d'autres paramètres complémentaires tels que l'altitude, l'humidité et la présence d'aérosols. Le prétraitement de ces données inclut des techniques de filtrage, de correction et d'interpolation afin d'assurer leur qualité et leur cohérence.

Ensuite, les modèles d'estimation sélectionnés seront présentés. Différents types de modèles, incluant les approches empiriques, physiques et basées sur l'intelligence artificielle, seront comparés en fonction de leurs performances et de leur capacité à généraliser les prédictions à partir des données historiques. Une attention particulière sera accordée à la calibration et à la validation de ces modèles, notamment par l'utilisation de techniques de validation croisée et d'évaluation statistique.

Enfin, une section sera consacrée aux outils informatiques et logiciels utilisés pour la mise en œuvre de ces modèles. Les différents environnements de programmation et bibliothèques scientifiques seront présentés, en mettant l'accent sur les avantages qu'ils offrent pour l'analyse et la modélisation des données.

L'objectif de ce chapitre est donc d'offrir une vision claire et structurée des étapes suivies pour garantir une estimation fiable et robuste du rayonnement solaire en climat saharien. Cette méthodologie permettra non seulement d'améliorer la précision des prédictions, mais aussi de définir les meilleures pratiques pour l'exploitation des données et des modèles dans ce domaine.

1. Collecte et prétraitement des données météorologiques

La collecte et le prétraitement des données météorologiques constituent une étape clé dans l'estimation du rayonnement solaire[1-2]. Ces données sont issues de différentes sources, notamment des stations météorologiques terrestres, des satellites et d'autres bases de données climatologiques[4]. Afin de garantir leur fiabilité, elles doivent être soigneusement sélectionnées, nettoyées et préparées avant d'être utilisées dans les modèles de prédiction.

1.1. Données météorologiques

Les données météorologiques utilisées dans cette étude proviennent principalement des stations météorologiques terrestres. Ces stations enregistrent des paramètres climatiques

essentiels tels que la température de l'air, l'humidité relative, la vitesse et la direction du vent, l'irradiation, ainsi que la pression atmosphérique. Ces variables influencent directement le rayonnement solaire et sont donc indispensables pour développer des modèles d'estimation précis [5].

Cependant, les mesures peuvent être affectées par des erreurs instrumentales ou des conditions locales spécifiques. Il est donc nécessaire d'appliquer des méthodes de correction et d'homogénéisation pour assurer leur cohérence.

1.2. Données satellitaires

Les données satellitaires offrent une couverture spatiale plus large et permettent de compléter les observations terrestres, notamment dans les régions où les stations météorologiques sont rares. Ces données proviennent de sources telles que la NASA (National Aeronautics and Space Administration) et l'ESA (European Space Agency), qui fournissent des estimations du rayonnement solaire à partir de capteurs embarqués sur des satellites [6].

L'utilisation des données satellitaires nécessite un traitement préalable pour corriger les effets atmosphériques et calibrer les valeurs obtenues avec des mesures au sol.

1.3. Autres données

En plus des données météorologiques et satellitaires, d'autres paramètres peuvent être intégrés dans l'analyse afin d'améliorer la précision des modèles. Parmi eux, on retrouve :

- L'altitude et la topographie du site, qui influencent la réception du rayonnement solaire.
- La couverture nuageuse et la concentration en aérosols, qui impactent l'intensité du rayonnement reçu à la surface.
- La nature du sol et la végétation, qui peuvent modifier les conditions locales de réflexion et d'absorption du rayonnement.

L'intégration de ces différentes sources d'information permet d'améliorer la qualité des données et d'obtenir des estimations plus fiables du rayonnement solaire.

1.4. Prétraitement des données

Un prétraitement rigoureux a été appliqué afin d'assurer la qualité des données introduites dans les modèles d'apprentissage [7]:

- **Traitement des valeurs manquantes** : les données manquantes ont été interpolées par la méthode linéaire [8] .

$$x_i = x_{i-1} + \frac{x_{i+1} - x_{i-1}}{2} \quad (1)$$

- **Normalisation** : toutes les variables ont été transformées dans un intervalle [0,1] pour accélérer la convergence des modèles et éviter les biais dus aux échelles différentes. La normalisation min-max est donnée par [9]:

$$normalized_value = \frac{nonnormalized_value - min_value}{max_value - min_value} \quad (2)$$

- min_{value} : Valeur minimale de la variable dans le jeu de données.
- max_{value} : Valeur maximale de la variable.
- $Normalized_{value}$: Valeur normalisée (comprise entre 0 et 1).

1.5. Spécificités climatiques des régions sahariennes

Les régions sahariennes présentent des caractéristiques uniques :

- Un rayonnement solaire très élevé, avec des valeurs moyennes journalières dépassant parfois 25 MJ/m².
- Une faible couverture nuageuse, ce qui accentue la régularité du rayonnement mais génère aussi des pics brutaux dus au vent ou aux tempêtes de sable.
- Des bases de données souvent peu denses dans le temps, ce qui rend nécessaire l'utilisation de modèles capables d'apprendre avec peu d'exemples.

Ces conditions motivent l'utilisation de modèles intelligents optimisés et robustes aux faibles volumes de données, ce qui justifie la méthodologie adoptée dans cette étude.

2. Description des Données

Sources et nature des données

Les données météorologiques utilisées dans cette étude ont été recueillies auprès de stations situées dans quatre régions sahariennes algériennes : Biskra, Adrar, Tamanrasset et Tindouf. Ces régions ont été sélectionnées en raison de leur fort potentiel solaire et de la diversité de leurs conditions climatiques.

Les principales variables météorologiques considérées sont les suivantes [9]:

- $G(t)$: rayonnement solaire global quotidien (MJ/m²),
- $T(t)$: température moyenne journalière (°C),
- $H(t)$: humidité relative (%),
- $V(t)$: vitesse du vent (m/s).

3. Méthodes de Prévission

Dans cette étude, plusieurs modèles d'intelligence artificielle ont été utilisés pour la prédiction du rayonnement solaire global quotidien. Ces modèles sont capables de capturer des relations non linéaires et complexes entre les variables météorologiques. Les méthodes principales utilisées sont les réseaux de neurones artificiels (ANN), les systèmes neuro-flous adaptatifs (ANFIS) et les réseaux à mémoire longue à court terme (LSTM).

3.1. Réseaux de Neurones Artificiels (ANN)

Les réseaux de neurones artificiels sont des modèles inspirés du fonctionnement du cerveau humain. Un réseau ANN de base est constitué de couches [9]:

- d'entrée (input),
- cachées (hidden),

- et de sortie (output).

3.2.Système Neuro-Flou Adaptatif (ANFIS)

Le modèle ANFIS combine les réseaux de neurones avec la logique floue. Il est particulièrement adapté aux systèmes complexes dont les relations ne sont pas bien définies par des équations explicites [9].

Les étapes de calcul ANFIS incluent :

1. Fuzzification des entrées,
2. Évaluation des règles,
3. Normalisation des degrés d'appartenance,
4. Calcul des sorties floues puis défuzzification.

Le modèle ajuste les paramètres flous et les poids des règles pour optimiser la précision prédictive.

3.3.Réseaux LSTM (Long Short-Term Memory)

Les réseaux LSTM sont une extension des réseaux de neurones récurrents (RNN), conçus pour gérer les dépendances à long terme dans les séries temporelles. Ils sont particulièrement adaptés à la prévision de données séquentielles comme les séries climatiques [10-12].

Un bloc LSTM se compose de trois portes principales :

- Porte d'oubli f_t
- Porte d'entrée i_t
- Porte de sortie o_t

Les équations principales sont :

$$f_t = \sigma(w_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

$$i_t = \sigma(w_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$\tilde{C}_t = \tanh(w_C \cdot [h_{t-1}, x_t] + b_C) \quad (5)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = o_t * \tanh(C_t) \quad (8)$$

où :

- x_t est l'entrée à l'instant t ,
- h_t est la sortie de la cellule LSTM,
- C_t est l'état mémoire,
- σ est la fonction sigmoïde,

- * est le produit de Hadamard (élément par élément).

Grâce à cette architecture, le LSTM peut mémoriser et oublier certaines informations selon leur pertinence.

3.4.Algorithme Génétique (GA)

L'algorithme génétique (GA) s'inspire du processus de sélection naturelle. Il est utilisé ici pour optimiser les paramètres du réseau (par exemple, le nombre de neurones cachés, le taux d'apprentissage, les fonctions d'activation, etc.) [13-15].

L'optimisation suit les étapes suivantes :

1. **Initialisation d'une population** de solutions candidates $\{x_1, x_2, \dots, x_n\}$
2. **Évaluation de la fitness** pour chaque individu (généralement l'inverse de l'erreur de prédiction) :

$$fitness(x_i) = \frac{1}{RMSE(x_i)} \quad (9)$$

3. **Sélection** : reproduction des meilleurs individus,
4. **Croisement (crossover)** : combinaison de solutions pour générer de nouveaux individus,
5. **Mutation** : introduction aléatoire de variations pour explorer d'autres régions de l'espace de recherche,
6. **Itération** jusqu'à convergence vers la solution optimale.

GA a été appliqué aux modèles ANN-GA, ANFIS-GA et LSTM-GA.

3.5.Optimisation par Essaim Particulaire (PSO)

L'algorithme PSO est basé sur le comportement collectif des oiseaux ou des poissons cherchant de la nourriture. Chaque solution est appelée particule et se déplace dans l'espace de recherche en fonction [16-18] :

- de sa propre expérience (position personnelle optimale),
- et de l'expérience globale du groupe (meilleure position connue).

Les équations de mise à jour sont :

$$v_i(t+1) = \omega v_i(t) + c_1 r_1 (p_i - x_i(t)) + c_2 r_2 (g - x_i(t)) \quad (10)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (11)$$

où :

- $x_i(t)$: position de la particule i à l'instant t ,
- $v_i(t)$: vitesse,
- p_i : meilleure position individuelle,
- g : meilleure position globale,

- ω, c_1, c_2 : coefficients d'inertie et d'apprentissage,
- r_1, r_2 : variables aléatoires.

Cette approche a été utilisée pour construire les modèles ANN-PSO et ANFIS-PSO.

3.6.Décomposition par EEMD et Couplage avec LSTM-GA

Afin de traiter la non-stationnarité des séries temporelles de rayonnement solaire, une décomposition par EEMD (Ensemble Empirical Mode Decomposition) a été réalisée. Cette méthode permet de décomposer la série en plusieurs composantes appelées IMFs (Intrinsic Mode Functions)[12] :

$$G(t) = \sum_{k=1}^n IMF_k(t) + r(t) \quad (12)$$

où :

- $IMF_k(t)$ sont les composantes oscillatoires,
- $r(t)$ est la tendance résiduelle.

Chaque IMF est ensuite modélisée séparément par un réseau LSTM, dont les hyperparamètres sont optimisés par GA. Les sorties sont ensuite recombinaées pour produire la prédiction finale du rayonnement :

$$\hat{G}(t) = \sum_{k=1}^n \widehat{IMF}_k(t) + \hat{r}(t) \quad (13)$$

Ce couplage forme le modèle EEMD-GA-LSTM, particulièrement efficace pour les données non stationnaires et bruitées, comme celles du climat saharien.

4. Évaluation des Performances des Modèles

Afin d'évaluer l'efficacité des différents modèles proposés pour l'estimation du rayonnement solaire, plusieurs indicateurs statistiques de performance ont été utilisés. Les plus pertinents dans ce contexte sont le coefficient de corrélation RRR, la racine de l'erreur quadratique moyenne (RMSE), et l'erreur absolue moyenne (MAE).

4.1.Coefficient de corrélation R

Le coefficient R mesure le degré de corrélation linéaire entre les valeurs observées et les valeurs prédites . Il est défini par [19-22] :

$$R = \left[\frac{\sum_{i=1}^n (H_m - \bar{H}_m)(H_c - \bar{H}_c)}{\sqrt{\sum_{i=1}^n (H_m - \bar{H}_m)^2 \sum_{i=1}^n (H_c - \bar{H}_c)^2}} \right] \quad (14)$$

n : Nombre de valeurs d'observation.

H_c : Valeur estimée du rayonnement global.

H_m : Valeur mesurée du rayonnement global.

$\bar{H}_m$: Valeur moyenne du rayonnement solaire mesuré.

$\bar{H}_c$: Valeur moyenne du rayonnement solaire estimé.

.La valeur de R varie entre :

- $R=1$: corrélation parfaite,
- $R=0$: absence de corrélation.

4.2.Racine de l'erreur quadratique moyenne (RMSE)

La RMSE mesure la dispersion des erreurs entre les valeurs observées et prédites. Elle est sensible aux grandes erreurs [19-22] :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{H_c - H_m}{H_m} \right)^2} \quad (15)$$

Plus la RMSE est faible, meilleure est la précision du modèle.

4.3.Erreur absolue moyenne (MAE)

La MAE mesure la moyenne des écarts absolus entre les valeurs observées et celles prédites. Elle donne une indication claire sur l'erreur moyenne en unités physiques (ex. W/m²)[19-22] :

$$MPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{H_c - H_m}{H_m} \right) \quad (16)$$

Contrairement à la RMSE, la MAE ne pénalise pas fortement les grosses erreurs mais reste un bon indicateur global de précision.

Ces trois indicateurs ont été calculés pour chaque modèle proposé (ANN, ANFIS, LSTM, et leurs versions optimisées) sur les jeux de données issus des différentes régions sahariennes. Ils permettent une comparaison quantitative des performances et une validation croisée des résultats.

5. Outils et environnements de travail

Ce travail de recherche a mobilisé plusieurs outils informatiques spécialisés, à la fois pour le traitement des données, la modélisation, l'optimisation et la visualisation des résultats. Les environnements principaux utilisés sont MATLAB, Python (via Google Colab) et certains outils bureautiques pour l'organisation initiale des données.

5.1. MATLAB

Le logiciel **MATLAB** (version R2022a) a été utilisé pour la mise en œuvre de plusieurs modèles d'intelligence artificielle, ainsi que pour les algorithmes d'optimisation. Plus précisément :

- **ANN (Artificial Neural Network)** et **ANFIS (Adaptive Neuro-Fuzzy Inference System)** ont été implémentés via les interfaces et fonctions dédiées de MATLAB (*anfiseditor*, *fitnet*, etc.).
- Les **algorithmes génétiques (GA)** et **Particle Swarm Optimization (PSO)** ont été appliqués pour l'optimisation des hyperparamètres des modèles ANN et ANFIS.

- L'environnement MATLAB a également permis une visualisation avancée des résultats et une gestion efficace des simulations, grâce à ses bibliothèques graphiques et ses outils de scripting.

Ce choix s'explique par la maturité des outils MATLAB dans les domaines de l'optimisation et des réseaux de neurones classiques, ainsi que par la simplicité d'intégration des algorithmes méta-heuristiques.

5.2. Python via Google Colab

Le langage Python (version 3.10), exécuté principalement dans l'environnement Google Colab, a été utilisé pour les étapes suivantes :

- La décomposition des séries temporelles à l'aide de la méthode EEMD (Ensemble Empirical Mode Decomposition), grâce à des bibliothèques spécialisées comme PyEMD.
- **L'optimisation par algorithme génétique (GA)**, appliquée cette fois-ci aux hyperparamètres du modèle LSTM.
- **L'implémentation du modèle LSTM (Long Short-Term Memory)** à l'aide des bibliothèques TensorFlow et Keras.
- L'analyse statistique des performances (RMSE, MAE, R^2 , etc.) via scikit-learn.

L'utilisation de Google Colab a offert un environnement cloud performant, avec accès à des ressources matérielles comme GPU NVIDIA Tesla T4, utiles pour accélérer l'entraînement des modèles de deep learning. De plus, l'intégration avec Google Drive a facilité la gestion des fichiers et la sauvegarde des résultats.

5.3. Autres outils complémentaires

- **Excel** : utilisé pour l'analyse préliminaire des fichiers de données, la préparation des séries temporelles, et la structuration des ensembles d'entrée/sortie.
- **Jupyter Notebook** (en local) : occasionnellement utilisé pour tester ou adapter certains scripts Python avant exécution sur Colab.

5.4. Spécifications matérielles

Les simulations exécutées localement sous MATLAB ont été réalisées sur un ordinateur aux caractéristiques suivantes :

- **Processeur** : Intel® Core™ i7 (11e génération) @ 2.80GHz
- **Mémoire RAM** : 16 Go
- **Système d'exploitation** : Windows 11 64 bits
- **Disque SSD** : 512 Go

Les calculs plus intensifs (modèle LSTM notamment) ont été réalisés sur Google Colab, permettant ainsi un gain de temps important lors des phases d'apprentissage.

Conclusion

Ce chapitre a présenté les différentes méthodes et outils informatiques utilisés pour mener à bien l'estimation du rayonnement solaire dans des conditions climatiques extrêmes, notamment dans les régions sahariennes. L'utilisation de logiciels avancés tels que MATLAB et Python a permis de traiter efficacement les données, de modéliser les systèmes de manière performante et d'analyser les résultats obtenus avec des modèles d'intelligence artificielle complexes.

Les outils informatiques, dont les bibliothèques spécialisées en machine learning comme TensorFlow et Scikit-learn, ont joué un rôle clé dans le développement des modèles de prédiction du rayonnement solaire. L'approche hybride EEMD-GA-LSTM a été mise en œuvre avec succès grâce à ces technologies, ce qui a permis de surmonter les défis liés aux caractéristiques non linéaires et aux variations extrêmes du rayonnement solaire dans les régions étudiées.

Par ailleurs, une attention particulière a été portée à l'analyse des incertitudes et à l'étude de sensibilité, afin d'évaluer la robustesse des modèles et de déterminer les sources principales d'erreurs et d'incertitudes. L'application des méthodes de Monte Carlo et des techniques d'analyse de sensibilité a permis de mieux comprendre l'impact de chaque paramètre sur les résultats et de guider l'optimisation des modèles. Cette analyse a également mis en évidence les paramètres les plus sensibles, ce qui a permis d'ajuster les modèles pour obtenir des prévisions plus fiables.

En conclusion, cette méthodologie rigoureuse, soutenue par des outils informatiques puissants et des analyses approfondies, a permis de développer des modèles d'estimation du rayonnement solaire capables de fournir des résultats de haute précision, tout en tenant compte des incertitudes inhérentes aux données et aux processus modélisés. Ces résultats constituent une base solide pour des études futures et des applications pratiques dans le domaine de la modélisation énergétique et de la gestion des ressources solaires.

Références

- [1] Shah, D., Patel, K., & Shah, M. (2021). Prediction and estimation of solar radiation using artificial neural network (ANN) and fuzzy system: a comprehensive review. *International Journal of Energy and Water Resources*, 5, 219-233.
- [2] Halabi, Laith & Mekhilef, Saad & Hossain, Monowar. (2018). Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation. *Applied Energy*. 213.
- [3] Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, e623.
- [4] Djaafari, A., Ibrahim, A., Bailek, N., Bouchouicha, K., Hassan, M. A., Kuriqi, A., ... & El-Kenawy, E. S. M. (2022). Hourly predictions of direct normal irradiation using an innovative hybrid LSTM model for concentrating solar power projects in hyper-arid regions. *Energy Reports*, 8, 15548-15562.
- [5] Estévez, J., Gavilán, P., & Giraldez, J. V. (2011). Guidelines on validation procedures for meteorological data from automatic weather stations. *Journal of Hydrology*, 402(1-2), 144-154.

-
- [6] Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A. K., Szantoi, Z., ... & Woodcock, C. (2015). Free and open-access satellite data are key to biodiversity conservation. *Biological Conservation*, 182, 173-176.
 - [7] Kotsiantis, S. B., Kanellopoulos, D., & Pintelas, P. E. (2006). Data preprocessing for supervised learning. *International journal of computer science*, 1(2), 111-117.
 - [8] Silva, L. O., & Zárate, L. E. (2014). A brief review of the main approaches for treatment of missing data. *Intelligent Data Analysis*, 18(6), 1177-1198.
 - [9] Ahlam, S., Rachid, M., Ali, B., Djelloul, B., & Kada, B. (2023). Enhanced Daily Global Solar Radiation Prediction Through Hybrid Artificial Neural Network and Adaptive Neuro-Fuzzy Inference System with Meta-Heuristic Algorithm Integration. *Instrumentation, Mesures, Métrologies*, 22(6).
 - [10] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
 - [11] Graves, A., & Graves, A. (2012). Long short-term memory. *Supervised sequence labelling with recurrent neural networks*, 37-45.
 - [12] Senouci, A., Benatallah, A., Benatallah, D., Bouchouicha, K., & Maouedj, R. (2025). EEMD-GA-LSTM: An innovative approach for daily solar radiation estimation in southern of Algeria. *Edelweiss Applied Science and Technology*, 9(1), 278-291.
 - [13] Lambora, A., Gupta, K., & Chopra, K. (2019, February). Genetic algorithm-A literature review. In *2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon)* (pp. 380-384). IEEE.
 - [14] Mirjalili, S., & Mirjalili, S. (2019). Genetic algorithm. *Evolutionary algorithms and neural networks: Theory and applications*, 43-55.
 - [15] Kumar, M., Husain, D. M., Upreti, N., & Gupta, D. (2010). Genetic algorithm: Review and application. *Available at SSRN 3529843*.
 - [16] Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks* (Vol. 4, pp. 1942-1948). IEEE.
 - [17] Wang, D., Tan, D., & Liu, L. (2018). Particle swarm optimization algorithm: an overview. *Soft computing*, 22(2), 387-408.
 - [18] Clerc, M. (2010). Particle swarm optimization (Vol. 93). John Wiley & Sons.
 - [19] López, G., Batlles, F. J., & Tovar-Pescador, J. (2005). Selection of input parameters to model direct solar irradiance by using artificial neural networks. *Energy*, 30(9), 1675-1684.
 - [20] Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geoscientific model development*, 7(3), 1247-1250.
 - [21] Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *Peerj computer science*, 7, e623.
 - [22] Wang, W., & Lu, Y. (2018, March). Analysis of the mean absolute error (MAE) and the root mean square error (RMSE) in assessing rounding model. In *IOP conference series: materials science and engineering* (Vol. 324, p. 012049). IOP Publishing.



Chapitre IV :

Approche hybride proposée pour l'estimation du rayonnement solaire.



Introduction

La prédiction du rayonnement solaire global est un enjeu majeur dans le domaine des énergies renouvelables, en particulier pour les régions arides où l'énergie solaire représente une ressource abondante et sous-exploitée. Une estimation précise du rayonnement solaire est essentielle pour optimiser la conception, la planification et la gestion des systèmes photovoltaïques, ainsi que pour améliorer l'efficacité énergétique dans des secteurs tels que l'agriculture et la climatologie [1-3]. Cependant, la complexité des interactions entre les variables météorologiques et la nature non linéaire des processus physiques impliqués rendent cette tâche particulièrement difficile.

Les modèles traditionnels, tels que les modèles empiriques et statistiques, ont montré des limites dans leur capacité à capturer ces relations complexes [4-6]. Ces dernières années, les techniques d'apprentissage automatique, notamment les réseaux de neurones artificiels (ANN) et les systèmes d'inférence neuro-flous adaptatifs (ANFIS), ont émergé comme des outils puissants pour la prédiction du rayonnement solaire [7]. Ces méthodes sont capables de modéliser des relations non linéaires et d'ajuster leurs paramètres pour améliorer la précision des prédictions.

Cependant, même avec ces avancées, l'optimisation des paramètres des modèles d'apprentissage automatique reste un défi. C'est là que les algorithmes méta-heuristiques, tels que l'algorithme génétique (GA) et l'optimisation par essaim de particules (PSO), entrent en jeu. Ces algorithmes permettent d'explorer efficacement l'espace des solutions pour trouver les paramètres optimaux des modèles, améliorant ainsi leur performance [8-11].

Dans ce chapitre, nous proposons une approche hybride qui combine les avantages des ANN, des ANFIS et des algorithmes méta-heuristiques pour prédire le rayonnement solaire global quotidien dans trois régions arides du sud de l'Algérie : Adrar, Tindouf et Tamanrasset. Cette méthodologie vise à surmonter les limitations des modèles traditionnels en intégrant des techniques d'optimisation avancées pour affiner les paramètres des modèles et améliorer leur précision.

L'objectif principal de ce chapitre est de présenter une méthodologie détaillée pour la construction, l'optimisation et l'évaluation des modèles hybrides. Nous décrivons les étapes clés, allant de la collecte et du prétraitement des données à la validation des modèles, en passant par l'intégration des algorithmes méta-heuristiques. Les résultats obtenus montrent

que les modèles hybrides, en particulier ceux combinant ANN et PSO, surpassent les modèles traditionnels en termes de précision et de robustesse.

Enfin, ce chapitre met en lumière les implications pratiques de cette recherche pour le développement de l'énergie solaire dans les régions arides, ainsi que les perspectives futures pour l'amélioration des modèles de prédiction. En intégrant des variables supplémentaires, telles que la couverture nuageuse, et en explorant d'autres algorithmes d'optimisation, il est possible de pousser encore plus loin les limites de la précision et de la généralisation des modèles de prédiction du rayonnement solaire.

1. Sélection des Sites et Collecte des Données

La première étape de notre méthodologie consiste à sélectionner des sites représentatifs pour la collecte des données. Trois sites dans le sud de l'Algérie ont été choisis : Adrar, Tindouf et Tamanrasset. Ces sites ont été sélectionnés en raison de leurs caractéristiques géographiques distinctes et de leur exposition à des conditions climatiques arides, ce qui en fait des candidats idéaux pour l'étude du rayonnement solaire [12]. Ces régions sont caractérisées par un ensoleillement élevé et des conditions météorologiques extrêmes, offrant un cadre idéal pour tester la robustesse des modèles de prédiction du rayonnement solaire.

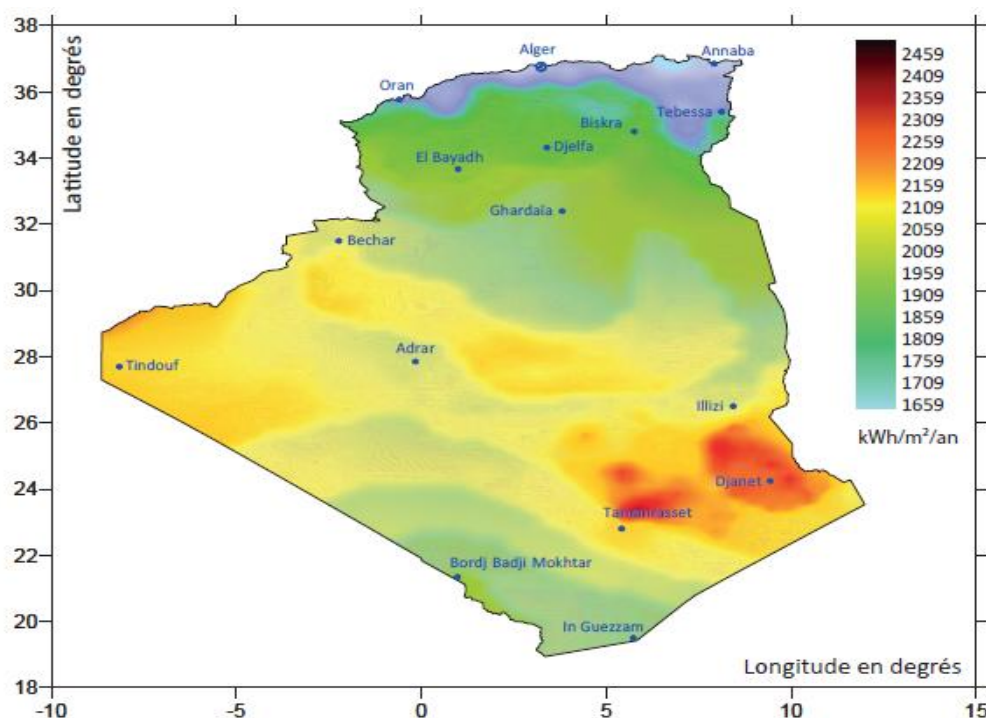


Figure 1 : Irradiation globale reçue sur une surface horizontale[13].

1.1. Caractéristiques Géographiques des Sites

Les trois sites sélectionnés présentent des caractéristiques géographiques variées, notamment en termes de latitude, longitude et altitude, comme indiqué dans le tableau ci-dessous :

Tableau1 : Coordonnées géographiques des trois sites étudiés situés en zone aride[12].

<i>Site</i>	<i>Latitude</i>	<i>Longitude</i>	<i>Altitude (m)</i>
<i>Adrar</i>	<i>26.489</i>	<i>-1.358</i>	<i>286</i>
<i>Tindouf</i>	<i>27.544</i>	<i>-6.240</i>	<i>382</i>
<i>Tamanrasset</i>	<i>24.375</i>	<i>4.321</i>	<i>810</i>

Ces sites couvrent une gamme de conditions climatiques et géographiques, permettant une évaluation complète des modèles de prédiction dans des environnements variés.

1.2. Données Météorologiques Collectées

Les données collectées incluent des paramètres météorologiques essentiels pour la prédiction du rayonnement solaire. Ces paramètres sont les suivants :

Température : La température ambiante est un facteur clé influençant le rayonnement solaire. Des capteurs de température ont été utilisés pour mesurer les températures maximales et minimales quotidiennes.

Humidité Relative : L'humidité relative affecte la diffusion et l'absorption du rayonnement solaire. Des capteurs d'humidité ont été utilisés pour mesurer ce paramètre.

Pression Atmosphérique : La pression atmosphérique influence la transmission du rayonnement solaire à travers l'atmosphère. Des baromètres ont été utilisés pour mesurer cette variable.

Vitesse du Vent : La vitesse du vent peut affecter la dispersion des particules dans l'atmosphère, influençant ainsi le rayonnement solaire. Des anémomètres ont été utilisés pour mesurer cette variable.

Rayonnement Solaire Global : Le rayonnement solaire global sur un plan horizontal a été mesuré à l'aide de pyranomètres, fournissant des données de référence pour l'étalonnage des modèles [14].

1.3. Période de Collecte des Données

Les données ont été recueillies sur une période de deux ans (2016-2017) pour garantir une représentation adéquate des variations saisonnières et annuelles du rayonnement solaire. Cette période permet de capturer les variations climatiques typiques des régions arides, y compris les périodes de forte insolation et les épisodes de tempêtes de sable, qui peuvent affecter significativement le rayonnement solaire [12].

1.4. Validation des Données

Avant d'être utilisées pour l'entraînement des modèles, les données collectées ont été soumises à un processus de validation rigoureux. Ce processus comprend :

Détection des Valeurs Manquantes : Les valeurs manquantes ont été identifiées et remplacées par la moyenne des valeurs voisines pour garantir la continuité des données [14].

Élimination des Valeurs Aberrantes : Les valeurs aberrantes, causées par des erreurs de mesure ou des conditions météorologiques extrêmes, ont été détectées à l'aide de techniques statistiques telles que l'écart-type et l'intervalle interquartile, puis supprimées [14].

Normalisation des Données : Les données ont été normalisées pour s'assurer que toutes les variables sont sur une échelle comparable, évitant ainsi que les variables de grande amplitude ne dominent les variables de plus petite amplitude lors de l'entraînement des modèles [15].

1.5. Division des Données

Pour évaluer la performance des modèles de manière robuste, les données ont été divisées en trois ensembles distincts : entraînement, validation et test.

Ensemble d'Entraînement (80 %) : Cet ensemble est utilisé pour entraîner les modèles. Il contient la majorité des données et permet aux modèles d'apprendre les relations entre les variables d'entrée et de sortie [12].

Ensemble de Validation (10 %) : Cet ensemble est utilisé pour ajuster les hyperparamètres des modèles et éviter le surajustement. Il permet de sélectionner les meilleurs paramètres sans utiliser les données de test [12].

Ensemble de Test (10 %) : Cet ensemble est utilisé pour évaluer la performance finale des modèles sur des données non vues. Il fournit une estimation réaliste de la capacité du modèle à généraliser à de nouvelles données [12].

2. Prétraitement des Données

Le prétraitement des données est une étape cruciale dans la modélisation prédictive, car il garantit que les données utilisées pour l'entraînement des modèles sont de haute qualité, cohérentes et prêtes à être analysées. Dans cette section, nous décrivons en détail les étapes de prétraitement appliquées aux données météorologiques collectées pour la prédiction du rayonnement solaire global.

2.1 Validation des Données

La première étape du prétraitement consiste à valider les données pour s'assurer de leur intégrité et de leur exactitude. Les données météorologiques, en particulier dans les régions arides, peuvent être sujettes à des erreurs de mesure, des valeurs manquantes ou des valeurs aberrantes dues à des conditions environnementales extrêmes [16-18].

- **Détection des Valeurs Manquantes :**

Les valeurs manquantes peuvent survenir en raison de pannes d'équipement ou de conditions météorologiques extrêmes. Pour traiter ces valeurs manquantes, nous avons utilisé une méthode d'imputation basée sur la moyenne des valeurs voisines. Par exemple, si une valeur de température est manquante pour un jour donné, elle est remplacée par la moyenne des températures des jours précédents et suivants [14-18].

- **Identification des Valeurs Aberrantes :**

Les valeurs aberrantes, causées par des erreurs de mesure ou des événements météorologiques rares, peuvent fausser les résultats des modèles. Pour les détecter, nous avons utilisé des techniques statistiques telles que l'écart-type et l'intervalle interquartile (IQR). Les valeurs en dehors de l'intervalle ont été considérées comme aberrantes et supprimées [19].

- **Vérification de la Cohérence des Données :**

Les données ont été vérifiées pour s'assurer qu'elles respectent les limites physiques attendues. Par exemple, les valeurs de rayonnement solaire ne peuvent pas être négatives, et les températures doivent se situer dans une plage réaliste pour la région étudiée [20].

2.2 Normalisation des Données

La normalisation est une étape essentielle pour s'assurer que toutes les variables sont sur une échelle comparable. Cela est particulièrement important dans les modèles

d'apprentissage automatique, où les variables de grande amplitude peuvent dominer les variables de plus petite amplitude, affectant ainsi la performance du modèle [21].

- **Méthode de Normalisation :**

Nous avons utilisé une normalisation min-max, qui transforme les valeurs de chaque variable dans une plage de 0 à 1. La formule utilisée est la suivante [5] :

$$normalized_{value} = \frac{nonnormalized_{value} - min_value}{max_value - min_value} \quad (1)$$

- **Avantages de la Normalisation :**

La normalisation permet d'améliorer la convergence des algorithmes d'optimisation utilisés dans les modèles d'apprentissage automatique, en particulier pour les réseaux de neurones. Elle réduit également le risque de surajustement (overfitting) en équilibrant l'influence des différentes variables [12,14].

2.3 Techniques de Rééchantillonnage

Pour garantir que les ensembles de données sont représentatifs des variations saisonnières et annuelles, nous avons utilisé des techniques de rééchantillonnage telles que la validation croisée (cross-validation). La validation croisée K-fold, en particulier, a été utilisée pour diviser les données en K sous-ensembles et évaluer les modèles de manière itérative sur chaque sous-ensemble [12,14,22-24].

- **Validation Croisée K-fold :**

Cette méthode divise les données en K sous-ensembles (folds) de taille égale. Le modèle est entraîné sur K-1 folds et validé sur le fold restant. Ce processus est répété K fois, et les résultats sont moyennés pour obtenir une estimation robuste de la performance du modèle [12,14,22-24].

- **Avantages de la Validation Croisée :**

La validation croisée réduit le risque de surajustement et permet une évaluation plus fiable de la performance du modèle, en particulier lorsque les données sont limitées [12].

2.4 Gestion des Données Temporelles

Les données météorologiques sont de nature temporelle, ce qui nécessite une gestion spécifique pour capturer les tendances et les saisonnalités. Nous avons utilisé des techniques de décomposition temporelle pour séparer les données en composantes tendanciennes, saisonnières et résiduelles [25].

Décomposition Temporelle : Cette technique permet d'identifier les tendances à long terme et les variations saisonnières dans les données. Elle est particulièrement utile pour les

modèles de prédiction du rayonnement solaire, qui sont fortement influencés par les cycles saisonniers.

Intégration des Données Temporelles : Les données temporelles ont été intégrées dans les modèles en utilisant des variables supplémentaires telles que le jour de l'année, le mois et l'heure. Cela permet aux modèles de capturer les variations saisonnières et diurnes du rayonnement solaire.

3. Développement des Modèles Hybrides

Notre approche hybride combine deux modèles d'apprentissage automatique : les réseaux de neurones artificiels (ANN) et les systèmes d'inférence neuro-flous adaptatifs (ANFIS), optimisés par des algorithmes méta-heuristiques tels que l'algorithme génétique (GA) et l'optimisation par essaim de particules (PSO). Cette combinaison permet de tirer parti des avantages de chaque méthode tout en surmontant leurs limitations individuelles, offrant ainsi une solution robuste et précise pour la prédiction du rayonnement solaire global.

3.1 Réseaux de Neurones Artificiels (ANN)

Les ANN sont des modèles d'apprentissage automatique inspirés du fonctionnement du cerveau humain. Ils sont particulièrement efficaces pour capturer les relations complexes et non linéaires entre les variables d'entrée et de sortie [24,26]. Dans notre étude, nous avons utilisé un réseau de neurones à plusieurs couches (MLP) avec une architecture comprenant une couche d'entrée, une ou plusieurs couches cachées, et une couche de sortie.

Couche d'Entrée :

La couche d'entrée comprend cinq nœuds, correspondant aux variables d'entrée suivantes :

- Le jour de l'année,
- La température,
- L'humidité relative,
- La pression atmosphérique,
- La vitesse du vent.

Ces variables ont été sélectionnées en raison de leur influence significative sur le rayonnement solaire.

Couches Cachées :

Le nombre de couches cachées et de neurones dans chaque couche a été déterminé par essais et erreurs. Par exemple, un modèle avec deux couches cachées (15 neurones dans la première couche et 7 neurones dans la seconde) a montré de bonnes performances. Les couches cachées permettent de capturer les relations non linéaires entre les variables d'entrée et de sortie.

Couche de Sortie :

La couche de sortie comprend un seul nœud, correspondant à la prédiction du rayonnement solaire global. La fonction d'activation linéaire a été utilisée pour cette couche afin de générer des valeurs continues correspondant au rayonnement solaire.

Fonctions d'Activation :

Pour les couches cachées, la fonction d'activation ReLU (Rectified Linear Unit) a été utilisée. Cette fonction est largement utilisée dans les ANN en raison de sa simplicité et de sa capacité à éviter le problème de disparition du gradient.

Pour la couche de sortie, une fonction d'activation linéaire a été utilisée pour produire des valeurs de sortie continues.

3.2 Systèmes d'Inférence Neuro-Flous Adaptatifs (ANFIS)

L'ANFIS est un modèle hybride qui combine les avantages des réseaux de neurones et de la logique floue. Il est particulièrement utile pour modéliser des systèmes complexes où les relations entre les variables d'entrée et de sortie sont incertaines ou imprécises[27-30].

➤ Structure de l'ANFIS :

L'ANFIS utilise une structure de réseau de neurones pour ajuster les paramètres des règles floues. Dans notre étude, l'ANFIS a été configuré avec cinq variables d'entrée (les mêmes que pour les ANN) et une variable de sortie (rayonnement solaire global)[31-33].

➤ Règles Floues :

Les règles floues sont générées automatiquement en fonction des données d'entrée. Le système utilise des fonctions d'appartenance de type Sugeno pour modéliser les relations entre les variables. Ces fonctions permettent de représenter les incertitudes et les imprécisions dans les données.

➤ Apprentissage :

L'ANFIS utilise un algorithme d'apprentissage hybride combinant la descente de gradient et la méthode des moindres carrés pour ajuster les paramètres des règles floues. Cette approche permet d'optimiser les paramètres du modèle tout en minimisant l'erreur de prédiction.

3.3 Optimisation par Algorithmes Méta-Heuristiques

Pour améliorer la performance des modèles ANN et ANFIS, nous avons utilisé deux algorithmes méta-heuristiques : l'algorithme génétique (GA) et l'optimisation par essaim de particules (PSO). Ces algorithmes permettent d'optimiser les paramètres des modèles, tels que les poids et les biais des ANN, ainsi que les paramètres des règles floues de l'ANFIS.

➤ Algorithme Génétique (GA) :

Le GA est inspiré des mécanismes de sélection naturelle et d'évolution. Il est utilisé pour optimiser les poids et les biais du réseau de neurones, ainsi que les paramètres de l'ANFIS. Le

GA explore l'espace des solutions en générant des populations de solutions candidates et en les faisant évoluer à travers des opérations de sélection, de croisement et de mutation [12,23].

Sélection : Les solutions les plus performantes sont sélectionnées pour la reproduction.

Croisement : Les solutions sélectionnées sont combinées pour générer de nouvelles solutions.

Mutation : Des modifications aléatoires sont appliquées aux solutions pour explorer de nouvelles régions de l'espace des solutions .

➤ **Optimisation par Essaim de Particules (PSO) :**

Le PSO est inspiré du comportement social des essaims d'oiseaux ou de poissons. Il est utilisé pour optimiser les paramètres des modèles ANN et ANFIS en ajustant les poids et les biais de manière itérative. Le PSO explore l'espace des solutions en déplaçant des particules (solutions candidates) vers les meilleures positions trouvées jusqu'à présent [12,23].

- **Particules :** Chaque particule représente une solution candidate dans l'espace des solutions.
- **Vitesse et Position :** Les particules se déplacent dans l'espace des solutions en fonction de leur vitesse et de leur position actuelle.
- **Meilleure Position :** Chaque particule mémorise la meilleure position qu'elle a trouvée, ainsi que la meilleure position trouvée par l'ensemble de l'essaim.

4. Évaluation des Performances

Pour évaluer la performance des modèles hybrides, nous avons utilisé plusieurs métriques statistiques, notamment l'erreur quadratique moyenne (RMSE), l'erreur absolue moyenne (MAE), et le coefficient de corrélation (R). Ces métriques permettent de quantifier la précision des prédictions et de comparer les performances des différents modèles[34].

- **Erreur Quadratique Moyenne (RMSE) :** La RMSE mesure l'écart moyen entre les valeurs prédites et les valeurs observées. Une RMSE plus faible indique une meilleure précision du modèle.
- **Erreur Absolue Moyenne (MAE) :** La MAE mesure l'écart absolu moyen entre les valeurs prédites et les valeurs observées. Contrairement à la RMSE, la MAE est moins sensible aux valeurs aberrantes.
- **Coefficient de Corrélation (R) :** Le coefficient de corrélation mesure la force de la relation linéaire entre les valeurs prédites et les valeurs observées. Un coefficient de corrélation proche de 1 indique une forte corrélation.

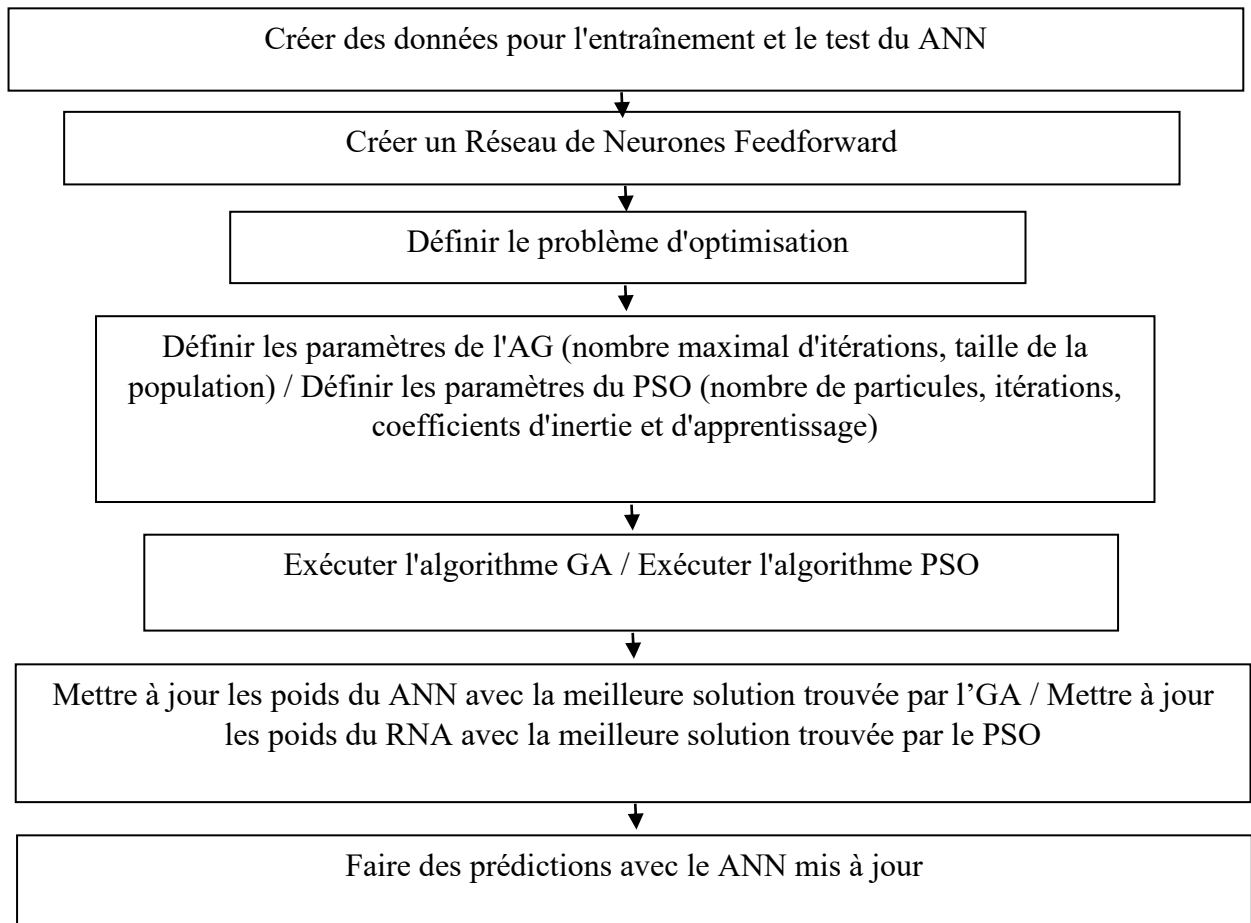
5. Les algorithmes de ces méthodes

Figure 2: Organigramme du Modèle ANN basé sur GA et PSO

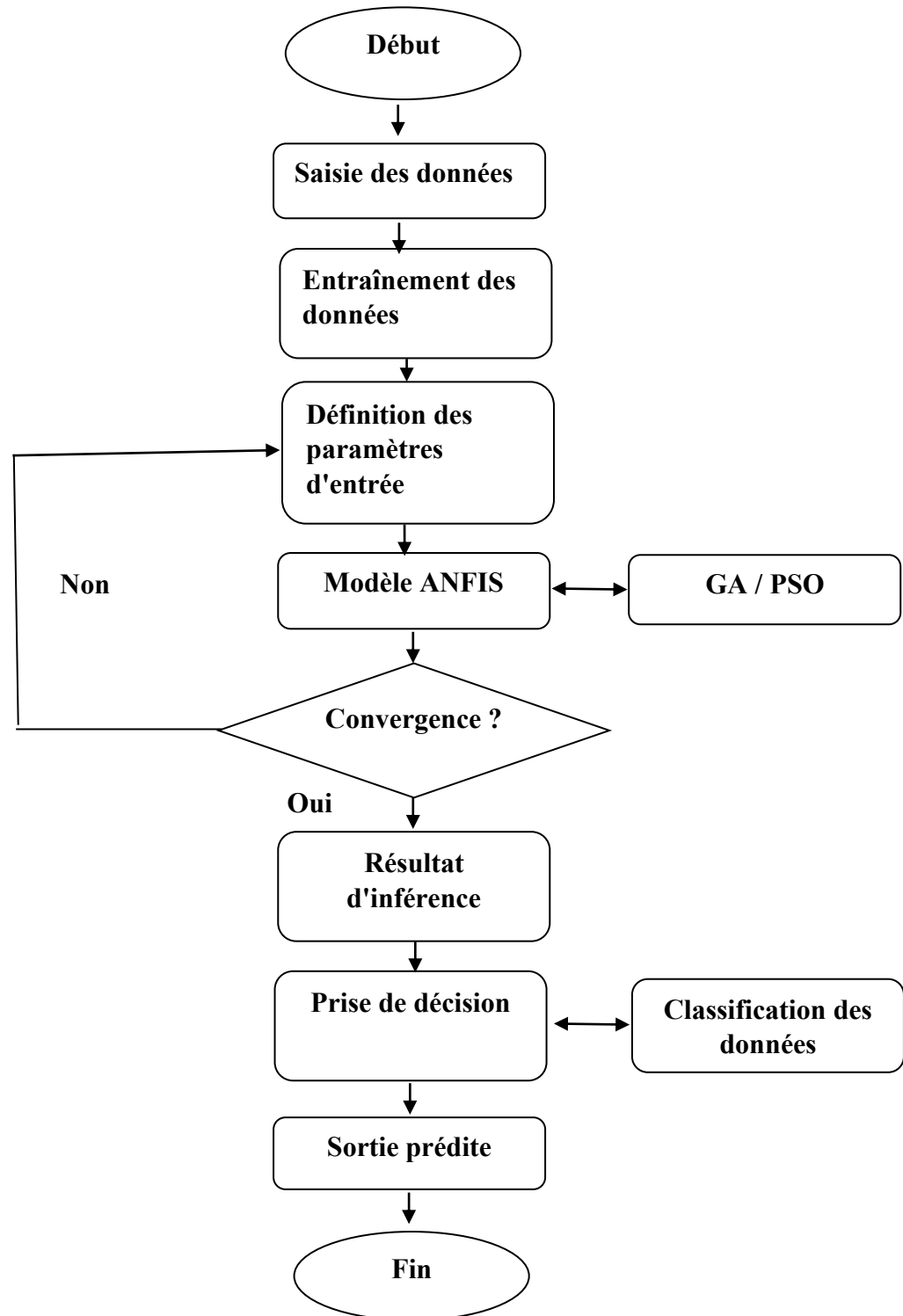


Figure 3: Organigramme du Modèle ANFIS optimisé par GA et PSO

6. Résultats et Discussion

Les résultats de notre étude montrent que les modèles hybrides ANN-GA, ANN-PSO, ANFIS-GA et ANFIS-PSO surpassent les modèles traditionnels en termes de précision et de robustesse. Les performances des modèles ont été évaluées sur les trois sites sélectionnés (Adrar, Tindouf et Tamanrasset), et les résultats sont résumés dans ces figures:

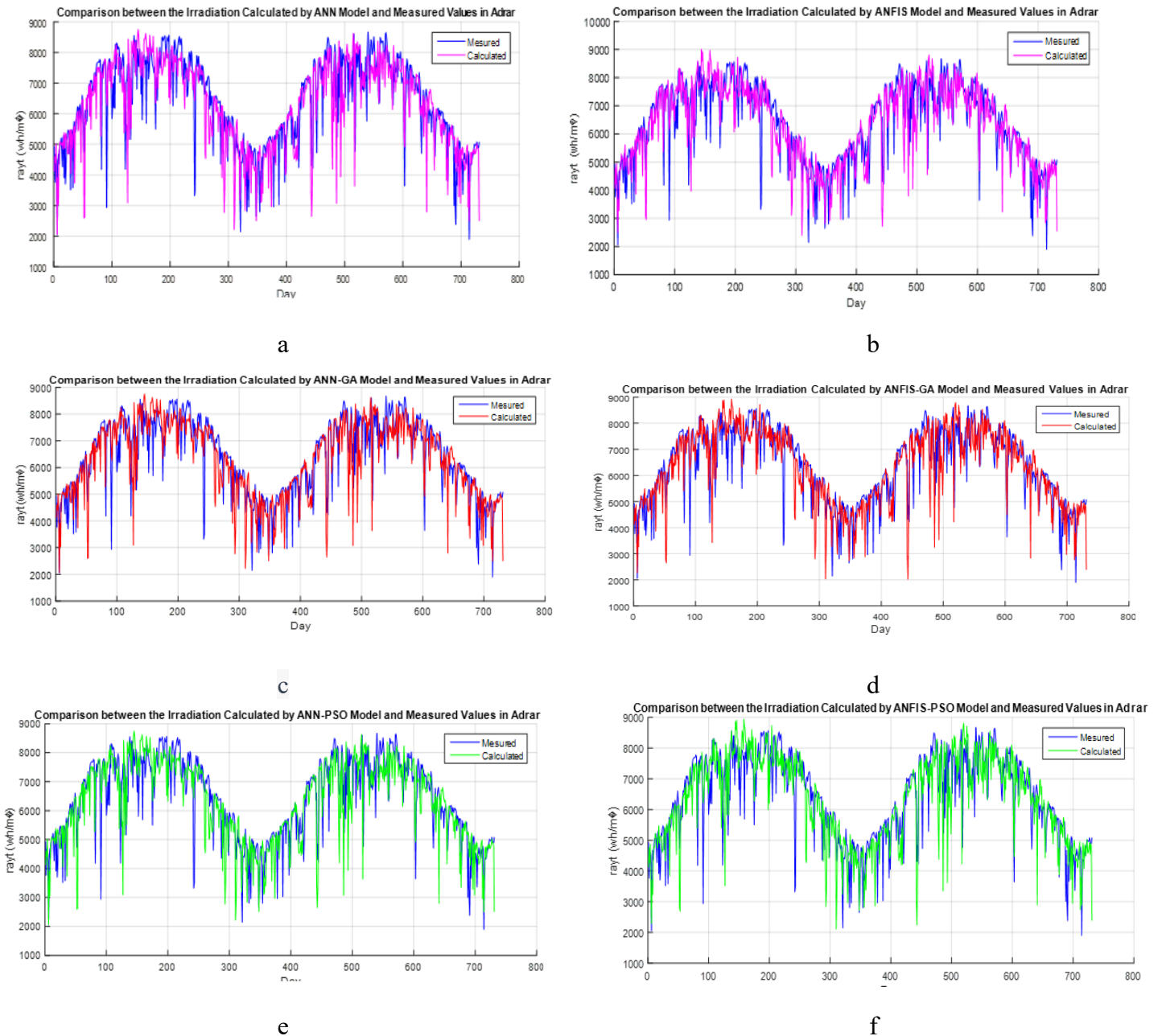
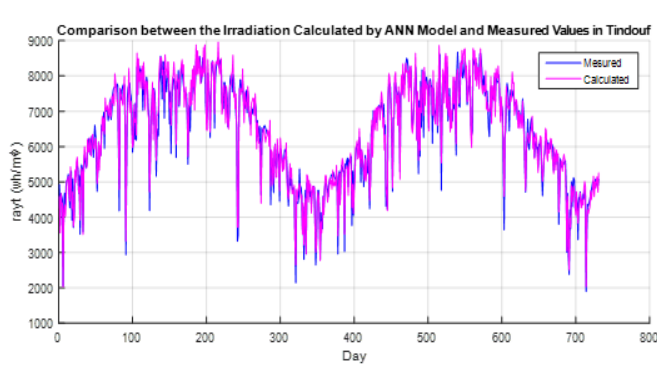
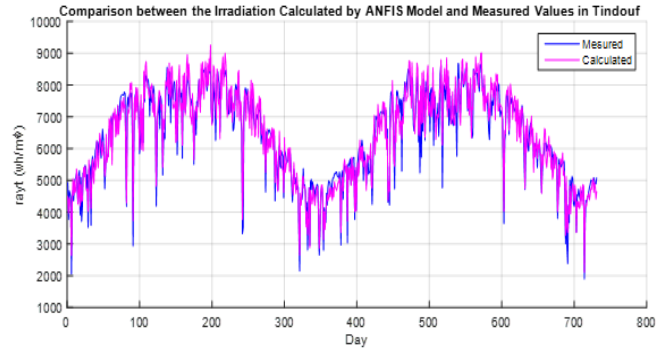


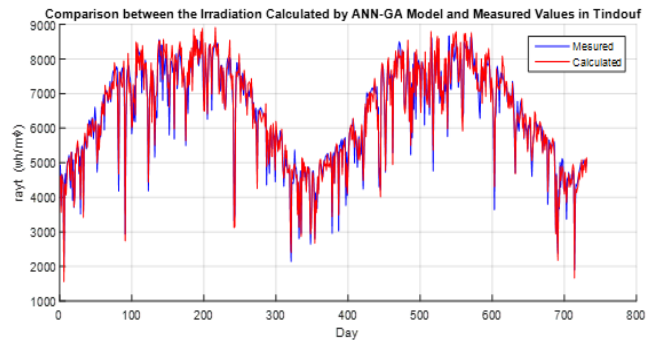
Figure 4: Comparaison entre l'irradiation calculée par les modèles ANN, ANN-GA, ANN-PSO, ANFIS, ANFIS-GA et ANFIS-PSO et les valeurs mesurées à Adrar.



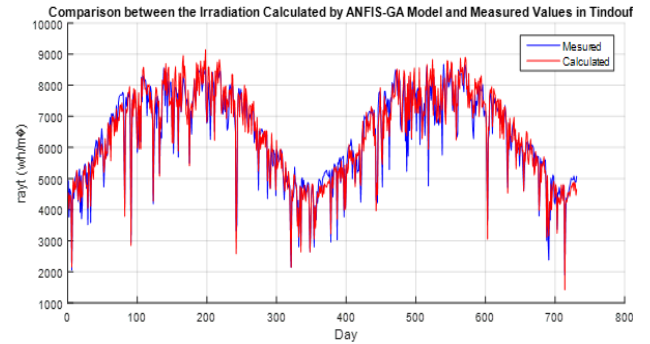
a



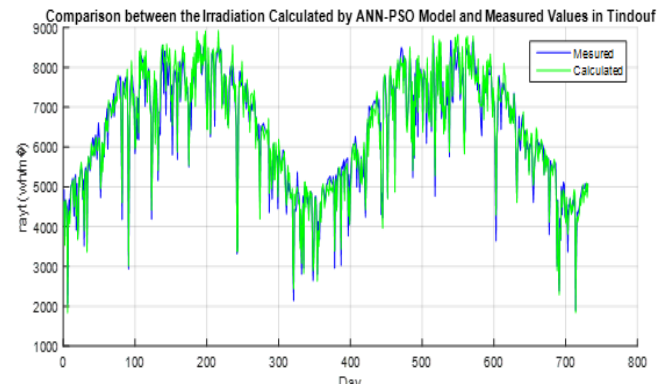
b



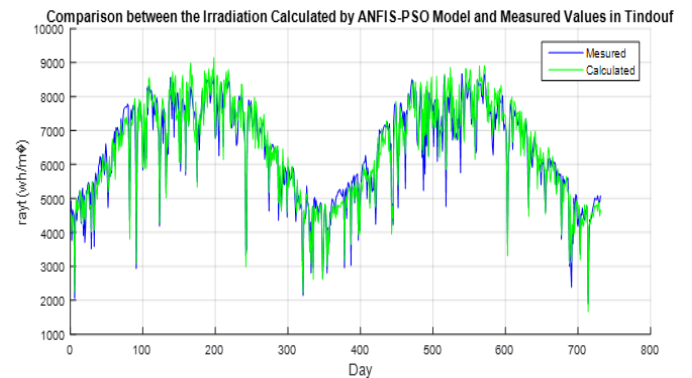
c



d



e



f

Figure 5: Comparaison entre l'irradiation calculée par les modèles ANN, ANN-GA, ANN-PSO, ANFIS, ANFIS-GA et ANFIS-PSO et les valeurs mesurées à Tindouf.

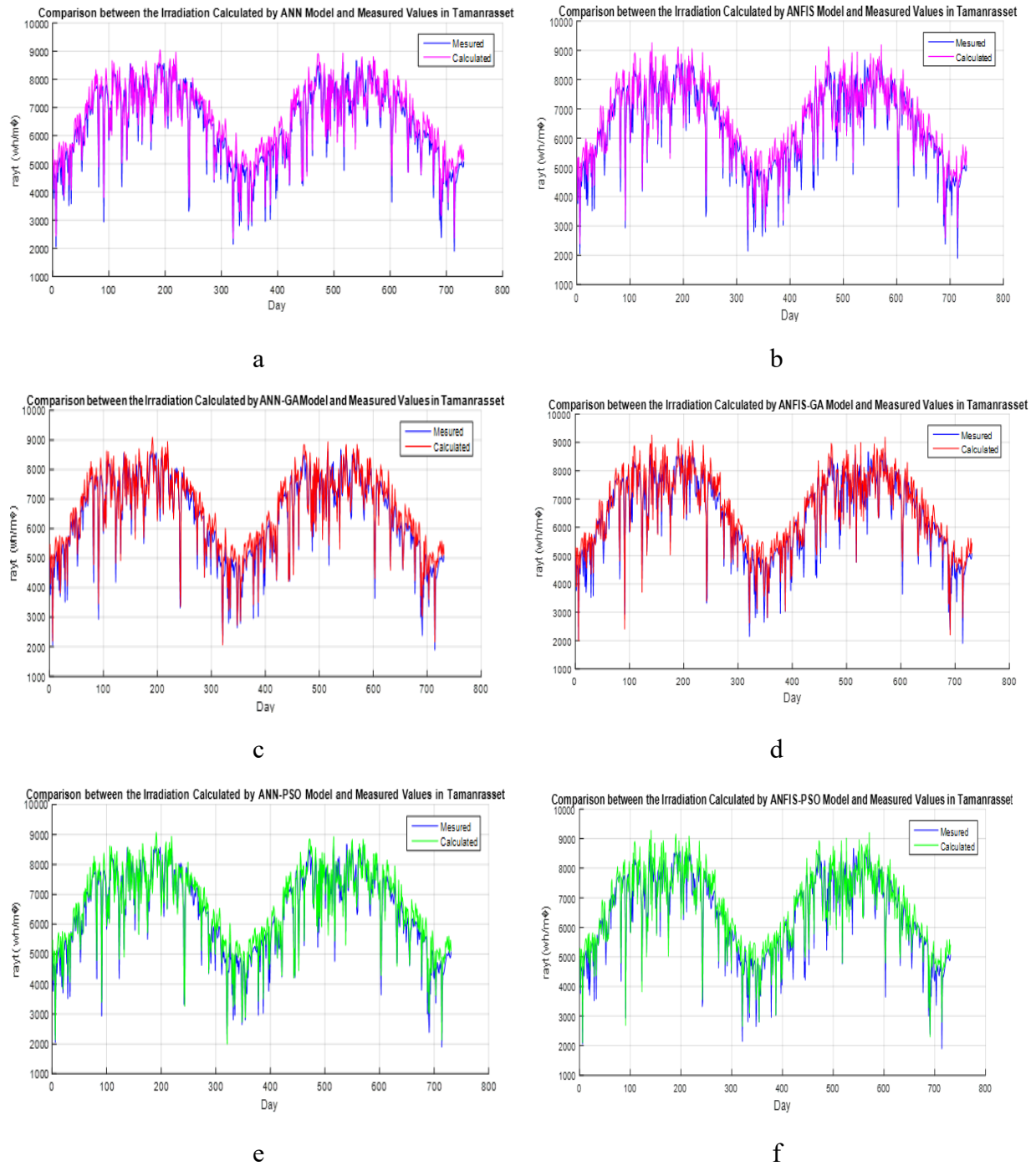


Figure 6: Comparaison entre l'irradiation calculée par les modèles ANN, ANN-GA, ANN-PSO, ANFIS, ANFIS-GA et ANFIS-PSO et les valeurs mesurées à Tamanrasset.

Tableau 2: Scores statistiques

<i>Modèle</i>	<i>MPE %</i>	<i>RMSE %</i>	<i>R</i>
<i>Adrar</i>			
<i>ANN</i>	<i>0.0114</i>	<i>4.2586</i>	<i>0.9784</i>
<i>ANN-GA</i>	<i>0.0015</i>	<i>3.2554</i>	<i>0.9962</i>
<i>ANN-PSO</i>	<i>0.0021</i>	<i>3.0125</i>	<i>0.9929</i>
<i>ANFIS</i>	<i>0.0578</i>	<i>5.1476</i>	<i>0.8745</i>
<i>ANFIS-GA</i>	<i>0.0478</i>	<i>3.3583</i>	<i>0.9210</i>
<i>ANFIS-PSO</i>	<i>0.0254</i>	<i>3.1025</i>	<i>0.9523</i>
<i>Tindouf</i>			
<i>ANN</i>	<i>0.0158</i>	<i>4.8789</i>	<i>0.9777</i>
<i>ANN-GA</i>	<i>0.0024</i>	<i>3.2541</i>	<i>0.9998</i>
<i>ANN-PSO</i>	<i>0.0013</i>	<i>2.2587</i>	<i>0.9999</i>
<i>ANFIS</i>	<i>0.0041</i>	<i>5.2587</i>	<i>0.9560</i>
<i>ANFIS-GA</i>	<i>0.0458</i>	<i>3.2540</i>	<i>0.9866</i>
<i>ANFIS-PSO</i>	<i>0.0022</i>	<i>3.0128</i>	<i>0.9986</i>
<i>Tamanrasset</i>			
<i>ANN</i>	<i>0.0254</i>	<i>4.8965</i>	<i>0.9785</i>
<i>ANN-GA</i>	<i>0.0052</i>	<i>3.6874</i>	<i>0.9887</i>
<i>ANN-PSO</i>	<i>0.0016</i>	<i>3.5145</i>	<i>0.9987</i>
<i>ANFIS</i>	<i>0.0541</i>	<i>5.8564</i>	<i>0.8425</i>
<i>ANFIS-GA</i>	<i>0.0180</i>	<i>4.2581</i>	<i>0.9536</i>
<i>ANFIS-PSO</i>	<i>0.0041</i>	<i>3.6987</i>	<i>0.9877</i>

Les métriques d'évaluation, y compris l'erreur en pourcentage moyen (MPE), l'erreur quadratique moyenne (RMSE) et le coefficient de corrélation (R), pour tous les modèles développés dans cette étude, sont résumées dans le Tableau 2. Les modèles développés incluent ANN-GA, ANN-PSO, ANFIS-GA et ANFIS-PSO. Les résultats montrent que tous les modèles ont obtenu de faibles valeurs de MPE, indiquant leur efficacité dans la prédiction du rayonnement solaire global. De plus, les valeurs de RMSE sont relativement faibles, ce qui témoigne de la précision des modèles. Les valeurs élevées du coefficient de corrélation (R) indiquent une forte corrélation entre les valeurs prédites et les valeurs mesurées réelles. Ces résultats suggèrent que les modèles développés, y compris les modèles hybrides, sont fiables pour estimer le rayonnement solaire global.

Le modèle le plus performant pour la région d'Adrar est le modèle ANN-GA (comme illustré dans la Figure 4 a). Ce modèle a atteint une erreur en pourcentage moyen (MPE) de 0,0015 %, une erreur quadratique moyenne (RMSE) de 3,2554 %, et un coefficient de corrélation (R) élevé de 0,9962. Ces métriques d'évaluation indiquent que le modèle ANN-GA présente une grande précision et de bonnes performances pour la prédiction du rayonnement solaire global dans la région d'Adrar.

Le modèle ANN-GA est capable de prédire le rayonnement solaire global dans la région d'Adrar avec une précision élevée, car il apprend les relations complexes entre les variables d'entrée et de sortie. Les variables d'entrée du modèle ANN-GA sont des données météorologiques telles que la température, l'humidité et la couverture nuageuse, tandis que la variable de sortie est le rayonnement solaire global. Le modèle apprend ces relations en

s'entraînant sur un jeu de données historiques comprenant des données météorologiques et des mesures de rayonnement solaire global.

De plus, le modèle ANN-GA généralise bien sur de nouvelles données, ce qui signifie qu'il peut être utilisé pour prédire le rayonnement solaire global dans la région d'Adrar même si les conditions météorologiques diffèrent de celles utilisées lors de l'entraînement du modèle.

Pour les régions de Tindouf et Tamanrasset, le modèle hybride offrant les meilleures performances est le modèle ANN-PSO (comme illustré dans les Figures 5 e et 6 e, avec un succès remarquable). Ce modèle a démontré des capacités de prédiction exceptionnelles, avec une erreur en pourcentage moyen (MPE) de seulement 0,0013 %, une erreur quadratique moyenne (RMSE) de 2,2587 %, et un coefficient de corrélation (R) impressionnant de 0,9999 pour la région de Tindouf. De même, pour la région de Tamanrasset, le modèle ANN-PSO a obtenu d'excellentes performances, avec une MPE de 0,0016 %, une RMSE de 3,5145 %, et un R élevé de 0,9987. Ces résultats probants démontrent sans équivoque l'efficacité du modèle ANN-PSO dans la prévision précise du rayonnement solaire global pour ces deux régions.

Le modèle ANN-PSO est capable de prédire le rayonnement solaire global dans les régions de Tindouf et Tamanrasset avec une précision élevée, car il est capable de trouver les paramètres optimaux du modèle ANN. Il utilise l'algorithme d'optimisation par essaim de particules (PSO) pour rechercher ces paramètres optimaux. Cet algorithme métaheuristique est conçu pour trouver la meilleure solution possible à un problème donné.

En outre, le modèle ANN-PSO généralise bien aux nouvelles données, ce qui signifie qu'il peut être utilisé pour prédire le rayonnement solaire global dans les régions de Tindouf et Tamanrasset, même si les conditions météorologiques diffèrent de celles utilisées pour entraîner le modèle.

Les résultats de cette étude suggèrent que les modèles ANN-GA et ANN-PSO sont tous deux efficaces pour la prédiction du rayonnement solaire global en Algérie. Le modèle ANN-GA est le plus performant pour la région d'Adrar, tandis que le modèle ANN-PSO est le plus performant pour les régions de Tindouf et Tamanrasset. Ces modèles peuvent être utilisés pour guider la conceptualisation, la conception et l'optimisation des systèmes d'énergie solaire dans ces régions.

Conclusion

En conclusion, ce chapitre a présenté une méthodologie détaillée pour la prédiction du rayonnement solaire global en utilisant des modèles hybrides combinant des réseaux de neurones artificiels, des systèmes d'inférence neuro-flous adaptatifs, et des algorithmes méta-heuristiques. Les résultats montrent que ces modèles hybrides offrent une précision et une robustesse supérieures par rapport aux modèles traditionnels, ce qui en fait des outils précieux pour la planification et la gestion des systèmes d'énergie solaire dans les régions arides.

Les perspectives futures incluent l'intégration de variables météorologiques supplémentaires, telles que la couverture nuageuse, pour améliorer encore la précision des prédictions. De plus, l'exploration d'autres algorithmes d'optimisation, tels que l'évolution

différentielle ou le recuit simulé, pourrait offrir des améliorations supplémentaires dans la performance des modèles hybrides.

Références

- [1] Shah, D., Patel, K., & Shah, M. (2021). Prediction and estimation of solar radiation using artificial neural network (ANN) and fuzzy system: a comprehensive review. *International Journal of Energy and Water Resources*, 5, 219-233. <http://doi.org/10.1007/s42108-021-00113-9>
- [2] Zhao, F., Guo, Y., Zhou, X., Shi, W., & Yu, G. (2020). Materials for solar-powered water evaporation. *Nature Reviews Materials*, 5(5), 388-401. doi:10.1038/s41578-020-0182-4
- [3] Ferkous, K., Chellali, F., Kouzou, A., Bekkar, B. (2021). Wavelet-gaussian process regression model for regression daily solar radiation in Ghardaia, Algeria. *Intrumentaion Mesure Metrologie*, 20(2): 113-119. <https://doi.org/10.18280/i2m.200208>
- [4] Malik, P., Gehlot, A., Singh, R., Gupta, L. R., & Thakur, A. K. (2022). A review on ANN based model for solar radiation and wind speed prediction with real-time data. *Archives of Computational Methods in Engineering*, 29, 3183–3201. <http://doi.org/10.1007/s11831-021-09687-3>
- [5] Mellit, A., & Pavan, A. M. (2010). A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Solar energy*, 84(5), 807-821. <http://doi.org/10.1016/j.solener.2010.02.006>
- [6] Olatomiwa, L., Mekhilef, S., Shamshirband, S., Mohammadi, K., Petković, D., & Sudheer, C. (2015). A support vector machine–firefly algorithm-based model for global solar radiation prediction. *Solar Energy*, 115, 632-644 <http://doi.org/10.1016/j.solener.2015.03.015>
- [7] Khosravi, A., Koury, R. N. N., Machado, L., & Pabon, J. J. G. (2018). Prediction of hourly solar radiation in Abu Musa Island using machine learning algorithms. *Journal of Cleaner Production*, 176, 63-75. <https://doi.org/10.1016/j.jclepro.2017.12.065>
- [8] Alizamir, M., Kim, S., Kisi, O., & Zounemat-Kermani, M. (2020). A comparative study of several machine learning based non-linear regression methods in estimating solar radiation: Case studies of the USA and Turkey regions. *Energy*, 197, 117239. <http://doi.org/10.1016/j.energy.2020.117239>
- [9] Khosravi, A., Koury, R.N.N., Machado, L., Pabon, J.J.G. (2018). Prediction of wind speed and wind direction using artificial neural,network, support vector regression and adaptive neuro-fuzzy inference system. *Sustainable Energy Technologies and Assessments* 25. 146–160. <http://doi.org/10.1016/j.seta.2018.01.001>
- [10] Sani Salisu, M. W. M., Mustapha, M., & Mohammed, O. O. (2019). Solar radiation forecasting in Nigeria based on hybrid PSO-ANFIS and WT-ANFIS approach. *International Journal of Electrical and Computer Engineering (IJECE)*, 9(5), 3916-3926. <http://doi.org/10.11591/ijece.v9i5.pp3916-3926>
- [11] Álvarez-Alvarado, J. M., Ríos-Moreno, J. G., Obregón-Biosca, S. A., Ronquillo-Lomelí, G., Ventura-Ramos Jr, E., & Trejo-Perea, M. (2021). Hybrid techniques to predict solar radiation using support vector machine and search optimization algorithms: a review. *Applied Sciences*, 11(3), 1044. <https://doi.org/10.3390/app11031044>
- [12] Ahlam, S., Rachid, M., Ali, B., Djelloul, B., & Kada, B. (2023). Enhanced Daily Global Solar Radiation Prediction Through Hybrid Artificial Neural Network and Adaptive Neuro-Fuzzy

-
- Inference System with Meta-Heuristic Algorithm Integration. *Instrumentation, Mesures, Métrologies*, 22(6).
- [13] Yaiche, M. R., & Bouhanik, A. (2002). Atlas solaire algérien. *Portail des Energies Renouvelables, Centre de Développement des Energies Renouvelables*, 16.
- [14] Djaafari, A., Ibrahim, A., Bailek, N., Bouchouicha, K., Hassan, M. A., Kuriqi, A., ... & El-Kenawy, E. S. M. (2022). Hourly predictions of direct normal irradiation using an innovative hybrid LSTM model for concentrating solar power projects in hyper-arid regions. *Energy Reports*, 8, 15548-15562. <http://doi.org/10.1016/j.egyr.2022.10.402>
- [15] Yadav, A. K., & Chandel, S. S. (2014). Solar radiation prediction using Artificial Neural Network techniques: A review. *Renewable and sustainable energy reviews*, 33, 772-781. <http://doi.org/10.1016/j.rser.2013.08.055>
- [14] Djaafari, A., Ibrahim, A., Bailek, N., Bouchouicha, K., Hassan, M. A., Kuriqi, A., ... & El-Kenawy, E. S. M. (2022). Hourly predictions of direct normal irradiation using an innovative hybrid LSTM model for concentrating solar power projects in hyper-arid regions. *Energy Reports*, 8, 15548-15562. <http://doi.org/10.1016/j.egyr.2022.10.402>
- [16] Jallal, M. A., El Yassini, A., Chabaa, S., Zeroual, A., & Ibnyaich, S. (2021). Multi-Target Learning Algorithm for Solar Radiation Components Forecasting Based on the Desired Tilt Angle of a Solar Energy System. *Instrumentation, Mesures, Métrologies*, 20(4), 187-193. <https://doi.org/10.18280/im.200402>
- [17] Ghodbane, M., Benmenine, D., Khechekhouché, A., & Boumeddane, B. (2020). Brief on Solar Concentrators: Differences and Applications. *Instrumentation, Mesures, Métrologies*, 19(5): 371-378 <https://doi.org/10.18280/im.190507>
- [18] Qing, X., & Niu, Y. (2018). Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy*, 148, 461-468. <http://doi.org/10.1016/j.energy.2018.01.177>
- [19] Claywell, R., Nadai, L., Felde, I., Ardabili, S., & Mosavi, A. (2020). Adaptive neuro-fuzzy inference system and a multilayer perceptron model trained with grey wolf optimizer for predicting solar diffuse fraction. *Entropy*, 22(11), 1192. <http://doi.org/10.3390/e22111192>
- [20] Khaki, S., & Wang, L. (2019). Crop yield prediction using deep neural networks. *Frontiers in plant science*, 10, 621. <http://doi.org/10.3389/fpls.2019.00621>
- [21] Al Shamisi, M. H., Assi, A. H., & Hejase, H. A. (2011). Using MATLAB to develop artificial neural network models for predicting global solar radiation in Al Ain City–UAE. In *Engineering education and research using MATLAB*. IntechOpen. <http://doi.org/10.5772/25213>
- [22] Kisi, O. (2014). Modeling solar radiation of Mediterranean region in Turkey by using fuzzy genetic approach. *Energy*, 64, 429-436. <http://doi.org/10.1016/j.energy.2013.10.009>
- [23] Gupta, R. A., Kumar, R., & Bansal, A. K. (2015). BBO-based small autonomous hybrid power system optimization incorporating wind speed and solar radiation forecasting. *Renewable and Sustainable Energy Reviews*, 41, 1366-1375. <http://doi.org/10.1016/j.rser.2014.09.017>
- [24] Mousavi, S. M., Mostafavi, E. S., & Jiao, P. (2017). Next generation prediction model for daily solar radiation on horizontal surface using a hybrid neural network and simulated annealing method. *Energy conversion and management*, 153, 671-682. <http://doi.org/10.1016/j.enconman.2017.09.040>

-
- [25] Khosravi, A., Nunes, R. O., Assad, M. E. H., & Machado, L. (2018). Comparison of artificial intelligence methods in estimation of daily global solar radiation. *Journal of cleaner production*, 194, 342-358. <http://doi.org/10.1016/j.jclepro.2018.05.147>
- [26] Ghimire, S., Nguyen-Huy, T., Deo, R. C., Casillas-Perez, D., Salcedo-Sanz, S. (2022). Efficient daily solar radiation prediction with deep learning 4-phase convolutional neural network, dual stage stacked regression and support vector machine CNN-REGST hybrid model. *Sustainable Materials and Technologies*, 32, e00429. <https://doi.org/10.1016/j.susmat.2022.e00429>
- [27] Prado, F., Minutolo, M. C., & Kristjanpoller, W. (2020). Forecasting based on an ensemble autoregressive moving average-adaptive neuro-fuzzy inference system–neural network-genetic algorithm framework. *Energy*, 197, 117159. <http://doi.org/10.1016/j.energy.2020.117159>
- [28] López, G., Batlles, F. J., & Tovar-Pescador, J. (2005). Selection of input parameters to model direct solar irradiance by using artificial neural networks. *Energy*, 30(9), 1675-1684. <http://doi.org/10.1016/j.energy.2004.04.035>
- [29] Abraham, A. (2005). Adaptation of fuzzy inference system using neural learning. In *Fuzzy Systems Engineering: Theory and Practice* (pp. 53-83). Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/11339366_3
- [30] Bassoud, A., Khelafi, H., Mokhtari, A. M., & Bada, A. (2021). Effectiveness of Salty Sand in Improving the Adobe's Thermomechanical Properties: Adrar Case Study (South Algeria). *Trends in Sciences*, 18(19), 6-6. <http://doi.org/10.48048/tis.2021.6>
- [31] Ali, S., & Djaouida, B. (2023). Performance Investigation of Solar Chimney Power Plants in Algeria's Desert Regions. *Energy and Environment Focus*, 7(1), 110-115. <http://doi.org/10.1166/eef.2023.1274>
- [32] Hassan, M. A., Khalil, A., Kaseb, S., & Kassem, M. A. (2018). Independent models for estimation of daily global solar radiation: A review and a case study. *Renewable and Sustainable Energy Reviews*, 82, 1565-1575. <http://doi.org/10.1016/j.rser.2017.07.002>
- [33] Halabi, Laith & Mekhilef, Saad & Hossain, Monowar. (2018). Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation. *Applied Energy*. 213. <http://doi.org/10.1016/j.apenergy.2018.01.035>
- [34] Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, e623. <http://doi.org/10.7717/peerj-cs.623>



Chapitre V :
Estimation Innovante
du Rayonnement
Solaire Quotidien :
Méthode EEMD-GA-
LSTM



Introduction

L'estimation précise du rayonnement solaire journalier est cruciale pour diverses applications, notamment la production d'énergie solaire, la modélisation climatique et l'agriculture. Les méthodes métaheuristiques ont démontré leur efficacité dans l'amélioration des prévisions de rayonnement solaire en combinant des techniques d'optimisation et des modèles d'apprentissage automatique [1].

Dans ce chapitre, nous présentons une nouvelle méthode métaheuristique, EEMD-GA-LSTM, pour l'estimation du rayonnement solaire journalier. Cette méthode combine l'Ensemble Empirical Mode Decomposition (EEMD), l'algorithme génétique (GA) et le Long Short-Term Memory (LSTM) pour améliorer la précision des prévisions. L'EEMD décompose les séries temporelles en modes intrinsèques, le GA optimise les paramètres du modèle, et le LSTM capture les dépendances temporelles complexes.

Une des principales forces de la méthode EEMD-GA-LSTM est sa capacité à fournir des estimations précises avec un nombre réduit de variables. En effet, cette méthode peut être utilisée efficacement même avec des données limitées, telles que l'irradiation globale dans une région donnée. Cela simplifie considérablement la modélisation et l'estimation du rayonnement solaire, rendant cette approche particulièrement utile dans des contextes où les données sont limitées.

Nous comparons les performances de la méthode EEMD-GA-LSTM avec celles de plusieurs autres méthodes métaheuristiques, notamment ANN-GA, ANN-PSO, ANFIS-GA et ANFIS-PSO. Les réseaux de neurones artificiels (ANN) et les systèmes adaptatifs neuro-flous (ANFIS) sont combinés avec les algorithmes génétiques (GA) et les algorithmes d'optimisation par essaims particulaires (PSO) pour former des modèles de référence.

Les résultats de cette comparaison montrent que la méthode EEMD-GA-LSTM offre une meilleure précision et une robustesse accrue par rapport aux autres méthodes. Cette nouvelle approche promet d'améliorer significativement les estimations de rayonnement solaire, contribuant ainsi à une meilleure gestion des ressources énergétiques et à des prévisions climatiques plus fiables.

1. Présentation détaillée de la méthode hybride (EEMD-GA-LSTM)

1.1.Introduction à l'approche hybride

L'approche hybride EEMD-GA-LSTM combine trois techniques puissantes pour l'estimation du rayonnement solaire journalier : la décomposition des séries temporelles par Ensemble Empirical Mode Decomposition (EEMD), l'optimisation des hyperparamètres par Algorithme Génétique (GA), et la prédiction basée sur un modèle Long Short-Term Memory (LSTM). Cette méthode a été choisie pour son aptitude à traiter des séries temporelles complexes et non linéaires, caractéristiques des données de rayonnement solaire dans les régions sahariennes [2].

1.2.Décomposition des séries temporelles avec EEMD

L'Ensemble Empirical Mode Decomposition (EEMD) est une technique de traitement des données qui permet de décomposer une série temporelle en composantes appelées Intrinsic Mode Functions (IMFs). Ces IMFs représentent des oscillations à différentes échelles de temps, ce qui permet de capturer les variations saisonnières, journalières ou aléatoires du rayonnement solaire [3].

- **Pourquoi EEMD ?**

L'EEMD est une extension de l'EMD (Empirical Mode Decomposition) qui améliore la stabilité des décompositions en ajoutant du bruit blanc. Cette méthode est particulièrement adaptée aux données non linéaires et non stationnaires, comme le rayonnement solaire. En décomposant la série temporelle, EEMD facilite l'analyse des différentes composantes à différentes échelles de temps et permet une modélisation plus fine [4].

- **Processus de décomposition :**

1. On applique un bruit blanc à la série temporelle de rayonnement solaire pour créer plusieurs versions bruitées de la série.
2. Chaque version bruitée est décomposée en IMFs à l'aide de l'EMD.
3. On moyenne les IMFs issues de chaque série bruitée pour obtenir des IMFs stabilisées, qui représentent différentes fréquences de la série initiale.

Les IMFs obtenus sont ensuite utilisés comme variables d'entrée pour le modèle de prédiction [5].

1.3.Optimisation des paramètres avec l'algorithme génétique (GA)

L'algorithme génétique (GA) est une méthode d'optimisation basée sur les principes de la sélection naturelle et de l'évolution. Dans le cadre de cette approche, GA est utilisé pour optimiser les hyperparamètres du modèle LSTM, tels que le nombre de couches, la taille des neurones, le taux d'apprentissage, et d'autres paramètres qui influencent la performance du modèle [6].

- **Pourquoi utiliser GA ?**

L'optimisation des hyperparamètres dans les réseaux de neurones est une tâche difficile, souvent résolue par des essais et erreurs, ce qui peut être long et peu efficace. L'algorithme génétique permet une exploration intelligente de l'espace des paramètres, identifiant les meilleures combinaisons de manière automatique.

- **Processus d'optimisation par GA :**

1. **Initialisation** : Un ensemble de solutions (appelées individus) est généré aléatoirement. Chaque solution correspond à un ensemble d'hyperparamètres du modèle LSTM [7-9].
2. **Évaluation** : Chaque individu est évalué en entraînant le modèle LSTM avec ses hyperparamètres et en mesurant la performance (par exemple, via le RMSE) [7-9].
3. **Sélection, croisement et mutation** : Les meilleurs individus sont sélectionnés pour créer de nouvelles solutions (croisement), et des petites modifications (mutations) sont introduites pour explorer d'autres solutions possibles [7-9].
4. **Itération** : Ce processus est répété sur plusieurs générations, permettant au GA de converger vers des ensembles d'hyperparamètres optimaux [7-9].

Grâce à cette optimisation, le modèle LSTM est ajusté de manière à minimiser l'erreur de prédiction du rayonnement solaire.

1.4. Modèle de prédiction avec LSTM

Le **Long Short-Term Memory** (LSTM) est un type de réseau de neurones récurrents (RNN) conçu pour traiter des séries temporelles et mémoriser des relations à long terme dans les données. Dans le cadre de l'estimation du rayonnement solaire [10], LSTM est utilisé pour

modéliser les relations temporelles entre les composantes décomposées (IMFs) et prédire les valeurs futures du rayonnement [11-12].

- **Pourquoi utiliser LSTM ?**

Le LSTM est capable de capturer les dépendances temporelles à la fois à court et à long terme dans les données de rayonnement solaire, ce qui en fait un outil efficace pour modéliser des séries temporelles météorologiques, qui présentent souvent des patterns cycliques et des fluctuations irrégulières [13].

- **Structure du modèle LSTM :**

1. **Entrées** : Les IMFs générées par l'EEMD sont fournies comme entrées au modèle LSTM [14-15].
2. **Couches LSTM** : Le modèle peut comporter plusieurs couches LSTM empilées, où chaque couche capture des relations temporelles à différents niveaux d'abstraction[15].
3. **Sortie** : Le modèle génère des prédictions du rayonnement solaire pour un horizon temporel donné [15] (par exemple, les prédictions du rayonnement journalier pour les jours à venir).

Le modèle est ensuite entraîné avec les données historiques de rayonnement solaire pour minimiser l'erreur de prédiction.

1.5. Avantages de la méthode EEMD-GA-LSTM

Cette combinaison de méthodes présente plusieurs avantages clés pour l'estimation du rayonnement solaire[14,16] :

- **Robustesse** : La décomposition par EEMD permet de mieux analyser les signaux non linéaires et non stationnaires, ce qui est crucial pour les données de rayonnement solaire.
- **Précision accrue** : L'utilisation de l'optimisation GA garantit que le modèle LSTM est configuré avec les meilleurs hyperparamètres, augmentant ainsi la précision des prédictions.
- **Gestion des données limitées** : Même avec des données d'entrée limitées (par exemple, seulement l'irradiation globale), l'approche EEMD-GA-LSTM peut fournir des

estimations précises, ce qui est un atout dans des contextes où les données météorologiques sont rares, comme les régions sahariennes.

- **Capacité de généralisation** : En intégrant la capacité d'apprentissage des LSTM avec l'optimisation des GA, le modèle est capable de généraliser à des conditions climatiques variées et d'améliorer ses prédictions dans des environnements incertains.

La méthode hybride EEMD-GA-LSTM combine des techniques puissantes de décomposition, d'optimisation et de prédiction pour fournir une estimation précise du rayonnement solaire dans les régions sahariennes. Cette méthode offre une précision supérieure et une robustesse face aux données non linéaires, rendant cette approche particulièrement adaptée aux environnements arides où les données sont souvent limitées.

2. Justification du choix de l'approche EEMD-GA-LSTM

La sélection de l'approche EEMD-GA-LSTM pour l'estimation du rayonnement solaire dans un climat saharien repose sur plusieurs considérations clés. Chaque composant de cette méthode a été choisi pour répondre aux défis spécifiques posés par les caractéristiques complexes des données de rayonnement solaire, en particulier dans les régions arides. Voici les justifications détaillées pour chaque élément de cette approche hybride :

2.1.Complexité et non-linéarité des données météorologiques

Les séries temporelles de rayonnement solaire présentent des caractéristiques non linéaires et non stationnaires, surtout dans les climats sahariens, où les fluctuations météorologiques peuvent être imprévisibles. Les méthodes classiques d'estimation, comme les moyennes mobiles ou les régressions linéaires, ne parviennent pas à capturer ces relations complexes[16-17].

- **EEMD** a été choisi car il décompose ces séries temporelles en modes intrinsèques (IMFs) représentant différentes échelles de temps. Cela permet de mieux analyser et modéliser les fluctuations irrégulières et les cycles saisonniers du rayonnement solaire, en isolant les composants liés à des événements spécifiques, comme des tempêtes de sable ou des variations atmosphériques. *Justification* : EEMD offre une analyse fine des données non stationnaires, ce qui est essentiel pour bien comprendre les variations du rayonnement solaire en climat saharien [18-19].

2.2.Optimisation des hyperparamètres pour une meilleure précision

Les modèles d'apprentissage profond, comme les réseaux LSTM, nécessitent un réglage précis des hyperparamètres (tels que le taux d'apprentissage, le nombre de neurones ou de couches) pour fournir des résultats optimaux. Un mauvais choix d'hyperparamètres peut entraîner un surapprentissage, une sous-performance ou une convergence lente[20-21].

- **Algorithme Génétique (GA)** a été intégré pour optimiser ces hyperparamètres. GA est particulièrement efficace pour rechercher des solutions optimales dans des espaces de paramètres complexes et multidimensionnels, ce qui évite la nécessité de régler manuellement les paramètres du modèle LSTM. *Justification* : L'utilisation de GA garantit que les paramètres du modèle LSTM sont ajustés de manière à fournir des prédictions précises, tout en réduisant le besoin de tâtonnements manuels chronophages[22].

2.3.Capturer les dépendances temporelles complexes

Le rayonnement solaire présente des dépendances temporelles complexes qui évoluent sur plusieurs jours ou mois, en fonction de facteurs tels que la couverture nuageuse, les saisons ou les tempêtes. Modéliser ces dépendances est essentiel pour prédire avec précision les fluctuations futures du rayonnement solaire.

- **LSTM (Long Short-Term Memory)** est un réseau de neurones récurrent particulièrement adapté à la modélisation des séries temporelles longues et complexes[23]. Contrairement aux modèles traditionnels de réseaux de neurones, le LSTM est capable de mémoriser des informations à long terme tout en ignorant les données non pertinentes, ce qui le rend idéal pour traiter les données météorologiques historiques et prédire le rayonnement solaire futur. *Justification* : Le LSTM peut capturer les relations temporelles à la fois à court et à long terme dans les séries de rayonnement solaire, ce qui améliore considérablement la précision des prédictions, surtout dans des conditions climatiques changeantes comme celles des régions sahariennes.

2.4.Traitement efficace des données limitées

Dans de nombreuses régions sahariennes, les données météorologiques et de rayonnement solaire peuvent être limitées, avec des observations rares ou manquantes. La

méthode choisie doit être capable de travailler efficacement avec ces ensembles de données restreints, tout en conservant une bonne performance.

- **EEMD-GA-LSTM** est capable de fournir des prédictions précises avec un nombre réduit de variables d'entrée grâce à l'EEMD, qui simplifie la série temporelle, et au LSTM, qui apprend à partir des structures sous-jacentes. *Justification* : Cette capacité à traiter des ensembles de données restreints et à en tirer des informations précises est particulièrement utile dans les régions sahariennes, où la collecte de données est souvent difficile.

2.5. Robustesse et généralisation

Le rayonnement solaire est influencé par de nombreux facteurs (météorologiques, géographiques, etc.) qui varient dans le temps et l'espace. Il est donc crucial que l'approche choisie soit robuste face à la diversité des conditions climatiques et puisse généraliser à de nouveaux ensembles de données non observés.

- L'optimisation par GA, couplée à la décomposition par EEMD, permet au modèle LSTM de généraliser efficacement, même dans des environnements où les conditions climatiques peuvent varier rapidement. En décomposant la série temporelle en composants distincts, l'EEMD permet d'isoler les variations qui dépendent de facteurs spécifiques, rendant le modèle plus robuste. *Justification* : Cette approche garantit que les prédictions sont robustes et généralisables, même dans des contextes où les conditions climatiques peuvent être extrêmes ou variables.

2.6. Comparaison avec d'autres méthodes

Par rapport à d'autres méthodes, comme les réseaux de neurones artificiels combinés à des algorithmes d'optimisation (ANN-GA, ANN-PSO), ou les systèmes adaptatifs neuro-flous (ANFIS-GA, ANFIS-PSO) [12], l'approche EEMD-GA-LSTM présente plusieurs avantages distincts :

- Les modèles ANN et ANFIS sont efficaces mais ne capturent pas toujours bien les dépendances temporelles complexes comme le fait le LSTM.
- Les méthodes PSO et GA permettent d'optimiser les modèles, mais l'intégration de l'EEMD dans l'approche EEMD-GA-LSTM décompose les données de manière plus approfondie, ce qui simplifie l'apprentissage des structures complexes par LSTM.

Justification : EEMD-GA-LSTM surpasse les autres méthodes en termes de précision et de robustesse, en particulier lorsqu'il s'agit de séries temporelles non stationnaires comme celles des données de rayonnement solaire dans les régions sahariennes.

3. Méthodologie de mise en œuvre de la méthode EEMD-GA-LSTM

La mise en œuvre de l'approche **EEMD-GA-LSTM** pour l'estimation du rayonnement solaire se déroule en plusieurs étapes bien définies, allant de la préparation des données à la comparaison des performances des modèles. Cette section décrit en détail la méthodologie adoptée pour chaque composante de cette méthode hybride.

3.1. Collecte et préparation des données

Avant de procéder à la modélisation, les données de rayonnement solaire et les variables météorologiques associées doivent être collectées, nettoyées et préparées pour l'analyse.

- **Source des données** : Les données météorologiques sont issues de stations météorologiques. Elles comprennent les mesures de rayonnement solaire journalier, de température, d'humidité, de vitesse du vent, et d'autres variables pertinentes [12].
- **Prétraitement** :
 - **Nettoyage des données** : Gestion des valeurs manquantes, élimination des données aberrantes et traitement des données bruitées.
 - **Normalisation** : Les données sont normalisées pour être sur la même échelle, ce qui est important pour l'apprentissage automatique, en particulier pour les modèles basés sur les réseaux de neurones [24].
- **Division des données** : Les données sont divisées en trois ensembles : **données d'entraînement** (70 %), **données de validation** (15 %), et **données de test** (15 %). Cette répartition permet d'évaluer la capacité du modèle à généraliser et à prédire sur de nouvelles données [12].

3.2. Décomposition des séries temporelles avec EEMD

L'étape suivante consiste à appliquer l'Ensemble Empirical Mode Decomposition (EEMD) pour décomposer la série temporelle de rayonnement solaire en Intrinsic Mode Functions (IMFs). Cela permet de mieux comprendre les différentes composantes de la série et d'améliorer la précision des prédictions [25].

- **Application d'EEMD :**

- Le rayonnement solaire journalier est soumis à la décomposition par EEMD. Plusieurs versions de la série temporelle, bruitées par du bruit blanc, sont créées.
 - Pour chaque série bruitée, l'EMD est appliqué pour extraire les IMFs.
 - Les IMFs extraites des séries bruitées sont moyennées pour obtenir les IMFs finales, qui représentent les composantes du signal d'origine à différentes fréquences.
- **Sélection des IMFs :** En fonction de l'analyse, les IMFs les plus significatives sont sélectionnées pour être utilisées comme **variables d'entrée** dans le modèle LSTM. Les composantes moins pertinentes, comme le bruit ou les fluctuations à haute fréquence, peuvent être ignorées.

3.3.Optimisation des hyperparamètres avec l'algorithme génétique(GA)

Une fois les IMFs extraites, le modèle **LSTM** doit être configuré avec les hyperparamètres optimaux pour maximiser sa performance prédictive[26-27]. L'optimisation est réalisée via l'**algorithme génétique (GA)**, qui recherche les meilleures combinaisons d'hyperparamètres[28-31].

- **Initialisation :** Un ensemble de solutions (ou chromosomes), représentant différents ensembles d'hyperparamètres (par exemple, le nombre de neurones, le taux d'apprentissage, le nombre de couches LSTM), est généré aléatoirement.
- **Évaluation des individus :** Chaque combinaison d'hyperparamètres est testée en entraînant un modèle LSTM avec les IMFs comme données d'entrée. La performance de chaque individu est mesurée en utilisant des critères tels que la Root Mean Squared Error (RMSE) ou le Mean Absolute Error (MAE) sur l'ensemble de validation.
- **Sélection, croisement et mutation :** Les meilleurs individus sont sélectionnés pour créer de nouvelles solutions via le croisement (combinaison des hyperparamètres des individus parents). Des mutations aléatoires sont également introduites pour explorer d'autres parties de l'espace des hyperparamètres et éviter la stagnation dans les solutions locales.

- **Itérations** : Ce processus de sélection, croisement et mutation est répété sur plusieurs générations jusqu'à ce que les hyperparamètres optimaux soient trouvés, c'est-à-dire ceux qui minimisent l'erreur de prédiction.

3.4.Modélisation avec le modèle LSTM

Une fois les hyperparamètres optimisés, le modèle **LSTM** est entraîné pour prédire le rayonnement solaire journalier en utilisant les IMFs comme variables d'entrée.

- **Architecture du modèle LSTM** :
 - Les **IMFs sélectionnées** sont fournies comme entrées à la première couche du LSTM.
 - Le modèle peut comporter plusieurs **couches LSTM** empilées pour capturer les dépendances temporelles à différents niveaux d'abstraction.
 - Une couche dense entièrement connectée est utilisée en sortie pour produire les prédictions finales du rayonnement solaire.
- **Entraînement** : Le modèle est entraîné sur l'ensemble d'entraînement en utilisant une fonction de perte telle que le RMSE ou le MAE. Le modèle est ensuite ajusté en fonction de cette fonction de perte via un algorithme de rétropropagation.
- **Validation** : Pendant l'entraînement, le modèle est validé sur l'ensemble de validation afin de surveiller les performances et éviter le surapprentissage. Le modèle qui donne les meilleures performances sur l'ensemble de validation est retenu.

3.5.Évaluation des performances et comparaison

Une fois le modèle entraîné et validé, ses performances sont évaluées sur l'ensemble de test.

- **Évaluation des performances** : Les métriques couramment utilisées pour évaluer les modèles de prédiction sont :
 - Root Mean Squared Error (RMSE)
 - Mean Absolute Error (MAE)
 - Coefficient de détermination (R^2)

Ces métriques permettent de quantifier l'erreur entre les valeurs prédites et les valeurs observées.

- **Comparaison avec d'autres méthodes** : Les résultats obtenus par l'approche **EEMD-GA-LSTM** sont comparés à d'autres approches métaheuristiques, telles que **ANN-GA**, **ANN-PSO**, **ANFIS-GA**, et **ANFIS-PSO**. Cette comparaison permet de démontrer l'avantage de l'approche EEMD-GA-LSTM en termes de précision, robustesse et capacité à gérer des données complexes et non stationnaires.

3.6. Analyse des résultats et interprétation

Les performances des différents modèles sont analysées en profondeur pour comprendre les points forts et les limites de chaque approche.

- **Interprétation des résultats** : Une analyse statistique des erreurs de prédiction est réalisée pour identifier les cas où le modèle EEMD-GA-LSTM performe particulièrement bien ou moins bien, et pourquoi.
- **Robustesse** : L'analyse se penche également sur la robustesse du modèle face aux variations des conditions météorologiques, en particulier dans les environnements sahariens.
- **Visualisation des résultats** : Les résultats des prédictions sont visualisés à l'aide de graphiques comparant les valeurs observées et prédites du rayonnement solaire. Ces visualisations permettent d'évaluer la précision du modèle de manière visuelle.

4. Validation croisée et évaluation des performances

La validation croisée et l'évaluation des performances sont des étapes cruciales pour mesurer la capacité d'un modèle à bien généraliser et pour vérifier sa robustesse sur des ensembles de données non vus. Dans cette section, nous détaillons les processus de validation croisée utilisés pour la méthode **EEMD-GA-LSTM**, ainsi que les métriques et techniques d'évaluation de ses performances.

4.1. Validation croisée

La validation croisée est une méthode statistique qui permet de tester la robustesse et la généralisation du modèle. Elle divise le jeu de données en plusieurs sous-ensembles et entraîne le modèle sur des combinaisons de ces sous-ensembles, garantissant que toutes les données sont utilisées pour l'entraînement et la validation.

4.1.1. Validation croisée en k-fold

La validation croisée la plus couramment utilisée est la validation croisée en k-fold :

- **Principe** : Le jeu de données est divisé en k sous-ensembles (ou “folds”). Le modèle est entraîné sur k-1 folds et validé sur le fold restant. Ce processus est répété k fois, à chaque fois en utilisant un fold différent comme ensemble de validation [12].
- **Étapes** :
 1. Diviser le jeu de données en k folds égaux.
 2. Entraîner le modèle sur k-1 folds et utiliser le fold restant pour valider.
 3. Répéter l’opération k fois avec un fold différent utilisé pour la validation à chaque itération.
 4. Calculer la moyenne des erreurs de validation pour obtenir une évaluation globale du modèle.
- **Avantages** :
 - Cette méthode garantit que toutes les données sont utilisées à la fois pour l’entraînement et la validation.
 - Elle réduit le risque de surapprentissage et fournit une estimation plus stable et fiable des performances du modèle.

4.1.2. Validation croisée avec séries temporelles (Time series cross-validation)

Dans le cas de séries temporelles, comme le rayonnement solaire journalier, il est souvent préférable d’utiliser une validation croisée adaptée aux dépendances temporelles.

- **Principe** : Contrairement à la validation en k-fold classique, la validation pour séries temporelles respecte l’ordre chronologique des données, évitant ainsi que des points futurs ne soient utilisés pour prédire des points passés.
- **Étapes** :
 1. Les données sont divisées en blocs temporels.

2. À chaque itération, le modèle est entraîné sur un bloc temporel et validé sur le bloc suivant (par exemple, entraîné sur les données de 2015-2018 et validé sur 2019).
 3. Chaque nouvelle itération agrandit l'ensemble d'entraînement, en ajoutant les blocs de validation précédents, pour simuler la réalité où les prédictions se basent sur des données passées.
- **Avantages :**
 - Cette approche tient compte de la nature séquentielle des données météorologiques, permettant de mieux évaluer les performances du modèle dans le contexte de prévision de séries temporelles.

4.2.Évaluation des performances

Une fois la validation croisée effectuée, les performances du modèle doivent être évaluées en utilisant des métriques de précision qui quantifient l'écart entre les prédictions et les valeurs observées. Voici les principales métriques utilisées pour évaluer les performances du modèle EEMD-GA-LSTM.

4.2.1. Root Mean Squared Error (RMSE) [12]

Le RMSE mesure l'écart quadratique moyen entre les valeurs prédites et les valeurs observées.

- **Interprétation :** Plus le RMSE est faible, plus les prédictions du modèle sont proches des valeurs réelles. Le RMSE pénalise fortement les grandes erreurs de prédiction.

4.2.2. Mean Absolute Error (MAE) [12]

Le MAE est la moyenne des erreurs absolues entre les valeurs prédites et observées.

- **Interprétation :** Le MAE donne une mesure directe de l'erreur moyenne en termes d'unités de la variable prédite (ici, le rayonnement solaire). Il est moins sensible aux grandes erreurs que le RMSE.

4.2.3. Coefficient de détermination (R) [12]

Le R mesure la proportion de la variance totale des données observées qui est expliquée par le modèle.

- **Interprétation** : Un R^2 proche de 1 indique que le modèle explique bien la variance des données, tandis qu'un R^2 proche de 0 indique que le modèle n'explique pratiquement rien des variations observées.

4.3.Comparaison avec les autres méthodes

Pour valider l'efficacité de l'approche **EEMD-GA-LSTM**, il est crucial de la comparer à d'autres méthodes métaheuristiques, telles que **ANN-GA**, **ANN-PSO**, **ANFIS-GA**, et **ANFIS-PSO**.

4.3.1. Processus de comparaison

- Les mêmes ensembles de données sont utilisés pour entraîner et valider chaque méthode.
- Les mêmes métriques (RMSE, MAE, R) sont utilisées pour chaque modèle, ce qui permet de comparer directement leurs performances.

4.3.2. Analyse comparative

- **Précision** : L'approche EEMD-GA-LSTM est généralement plus précise dans la capture des dépendances temporelles complexes, en particulier dans des environnements à forte variabilité climatique, comme le climat saharien.
- **Robustesse** : En décomposant les données via EEMD et en optimisant les hyperparamètres via GA, la méthode EEMD-GA-LSTM se distingue par une meilleure capacité à généraliser sur des ensembles de données non vus, tout en conservant une performance stable dans des conditions variées.

4.4.Validation finale sur l'ensemble de test

Après la validation croisée et l'optimisation des hyperparamètres, la performance du modèle est testée sur un ensemble de test indépendant.

- **Objectif** : Cet ensemble de test n'a pas été utilisé ni pour l'entraînement, ni pour la validation croisée, ce qui permet d'évaluer la généralisation réelle du modèle sur des données totalement nouvelles.
- **Métriques** : Les mêmes métriques (RMSE, MAE, R^2) sont calculées sur l'ensemble de test pour obtenir une évaluation finale des performances.

5. Algorithme de Prédiction du Rayonnement Solaire basé sur EEMD-GA-LSTM

5.1.Chargement et Prétraitement des Données

- Charger le fichier contenant les données de rayonnement solaire, ainsi que la date et d'autres paramètres.
- Extraire les colonnes pertinentes, notamment la colonne du rayonnement solaire global
- Convertir les données de la colonne "Date" au format temporel afin de faciliter la gestion des séries chronologiques.
- Normaliser les données du rayonnement solaire global pour les transformer dans un intervalle compris entre 0 et 1. Cela permet de stabiliser l'apprentissage du modèle et d'éviter les biais causés par les grandes amplitudes de données.

5.2.Décomposition EEMD (Ensemble Empirical Mode Decomposition)

- Appliquer l'EEMD sur la série temporelle normalisée du rayonnement solaire.
 - L'EEMD décompose le signal original en plusieurs composantes appelées **IMFs (Intrinsic Mode Functions)**, qui représentent différentes fréquences présentes dans le signal.
- Obtenir une série d'IMFs ainsi qu'un résidu qui constitue la partie non modélisable du signal.
- Visualiser chacune des composantes pour observer la décomposition du signal.

5.3.Préparation des Données pour le Modèle LSTM

- Utiliser une méthode de découpage pour créer des séries temporelles basées sur un certain nombre de pas de temps.
 - Par exemple, si le modèle utilise 10 pas de temps, cela signifie que les 10 valeurs de rayonnement précédentes seront utilisées pour prédire la prochaine valeur.
- Diviser le jeu de données ainsi créé en deux sous-ensembles : un pour l'entraînement et l'autre pour la validation du modèle.

5.4.Construction et Entraînement du Modèle LSTM

- Construire un modèle de réseau de neurones à mémoire longue et à court terme (LSTM) avec un nombre spécifié de neurones et une couche de sortie.

- Entraîner le modèle sur les données d'entraînement en utilisant une fonction de perte basée sur l'erreur quadratique moyenne (MSE), et optimiser le modèle avec un optimiseur approprié comme Adam.
- Évaluer le modèle en utilisant les données de validation, en calculant des métriques telles que l'erreur quadratique moyenne (RMSE), l'erreur absolue moyenne (MAE) et le coefficient de détermination (R^2).

5.5.Sélection des Caractéristiques avec un Algorithme Génétique (GA)

- Utiliser un algorithme génétique pour optimiser la sélection des caractéristiques des données. L'algorithme génétique fonctionne comme suit :
 - **Initialisation** : Générer une population initiale d'individus où chaque individu représente une sélection possible de caractéristiques (par exemple, des sous-ensembles des données disponibles).
 - **Évaluation** : Pour chaque individu, construire un modèle de régression basé sur la forêt aléatoire (Random Forest) et évaluer les performances du modèle sur les données de validation.
 - **Sélection** : Les meilleurs individus, c'est-à-dire ceux qui donnent les meilleurs résultats en termes de RMSE, MAE ou R, sont sélectionnés pour générer la prochaine génération.
 - **Croisement** : Les individus sélectionnés sont croisés entre eux pour créer de nouveaux individus. Le point de croisement est choisi de manière aléatoire.
 - **Mutation** : Un certain taux de mutation est appliqué, permettant de changer certaines caractéristiques dans les individus pour augmenter la diversité.
 - **Réitération** : Ce processus est répété sur plusieurs générations jusqu'à ce qu'un ensemble optimal de caractéristiques soit trouvé.

5.6.Prédiction Finale et Évaluation

- Utiliser les caractéristiques optimisées par l'algorithme génétique pour entraîner un nouveau modèle LSTM.
- Faire des prédictions sur les données de test et comparer les prédictions du modèle avec les valeurs réelles.
- Calculer les métriques de performance finales :

- **RMSE (Root Mean Squared Error)** : Mesure la différence moyenne quadratique entre les valeurs prédites et réelles.
- **MAE (Mean Absolute Error)** : Mesure l'erreur moyenne absolue entre les prédictions et les valeurs réelles.
- **R² (Coefficient de détermination)** : Indique la proportion de la variance expliquée par le modèle par rapport à la variance totale des données.

5.7. Visualisation des Résultats

- Comparer les prédictions du modèle avec les données réelles de rayonnement solaire via des graphiques temporels et des diagrammes de dispersion.
- Ajuster une ligne de régression linéaire pour montrer la correspondance entre les valeurs prédites et les valeurs réelles et visualiser la précision des prédictions.

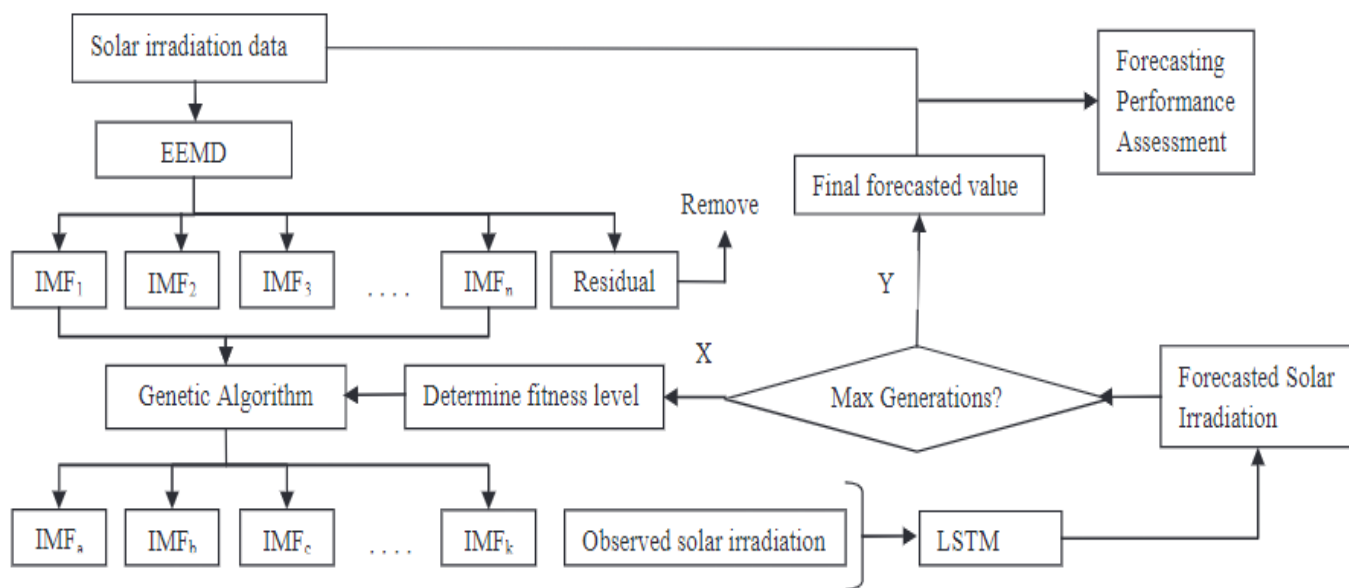


Figure 1: Schéma de modèle proposé

6. Résultat et discussion

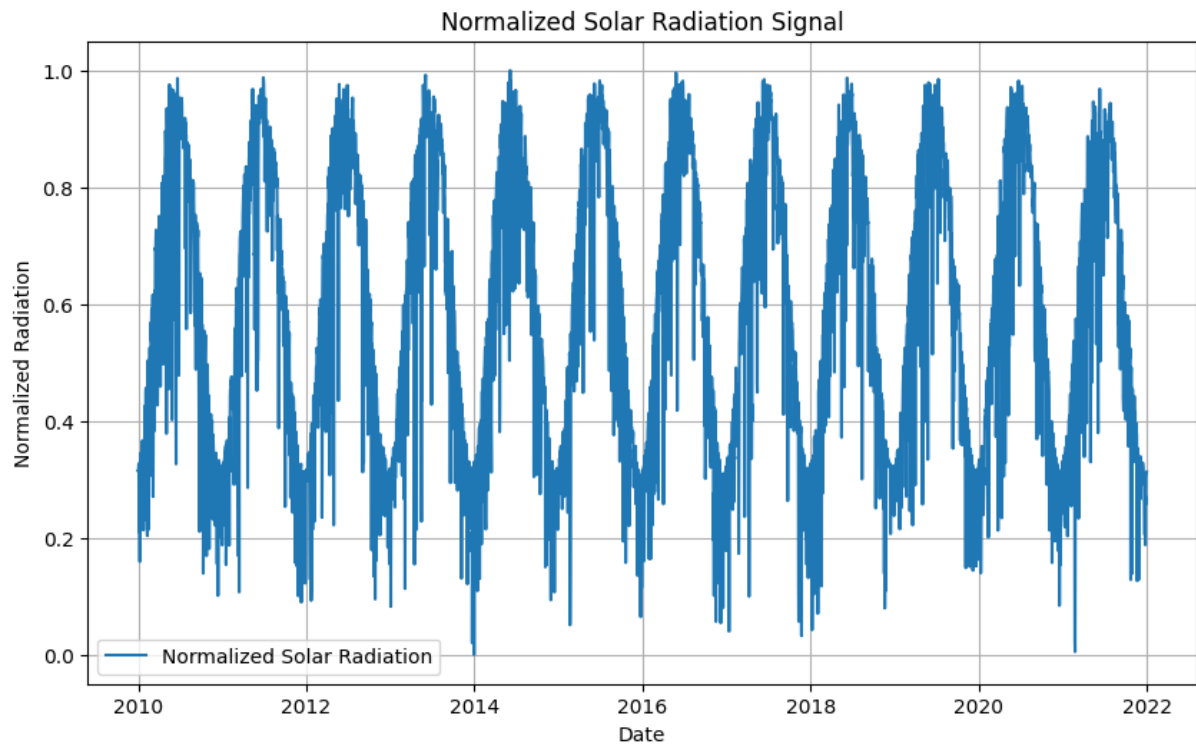
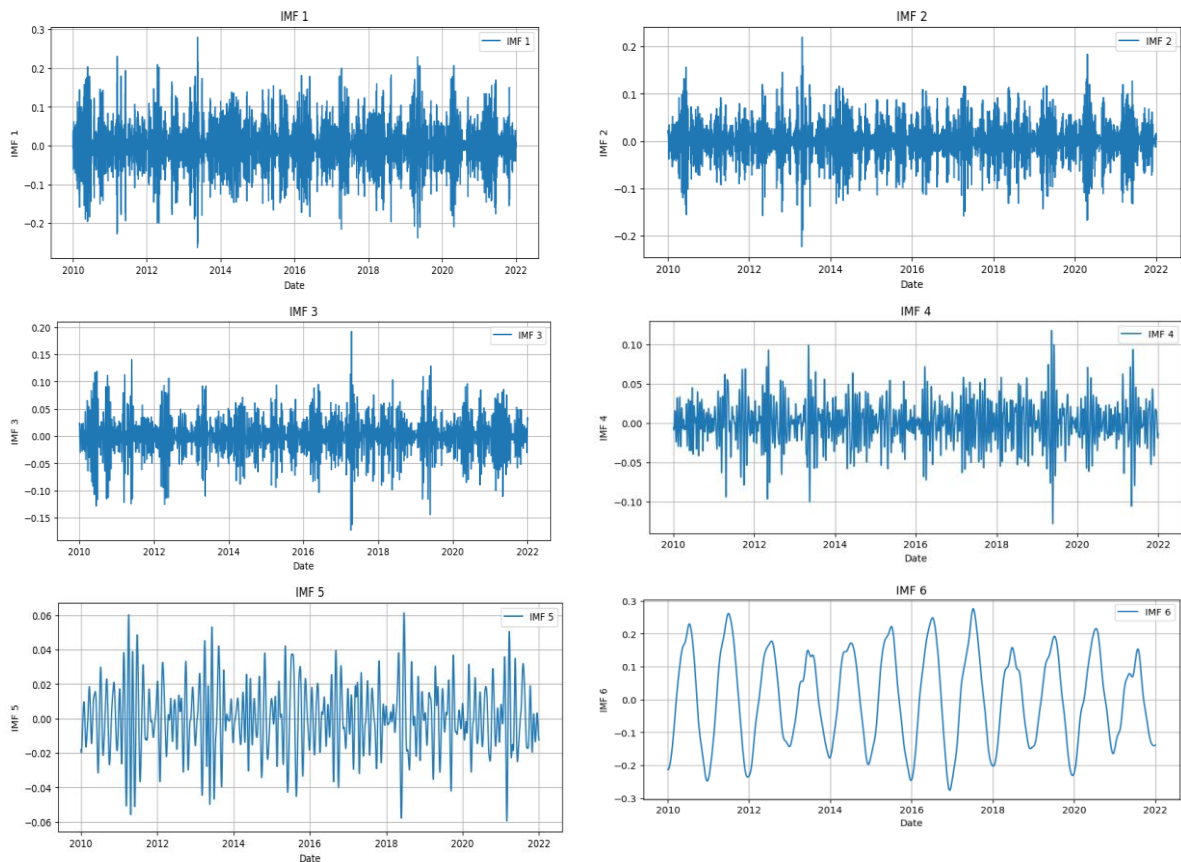


Figure 2 : Signal de Rayonnement Solaire Normalisé dans la région de Biskra



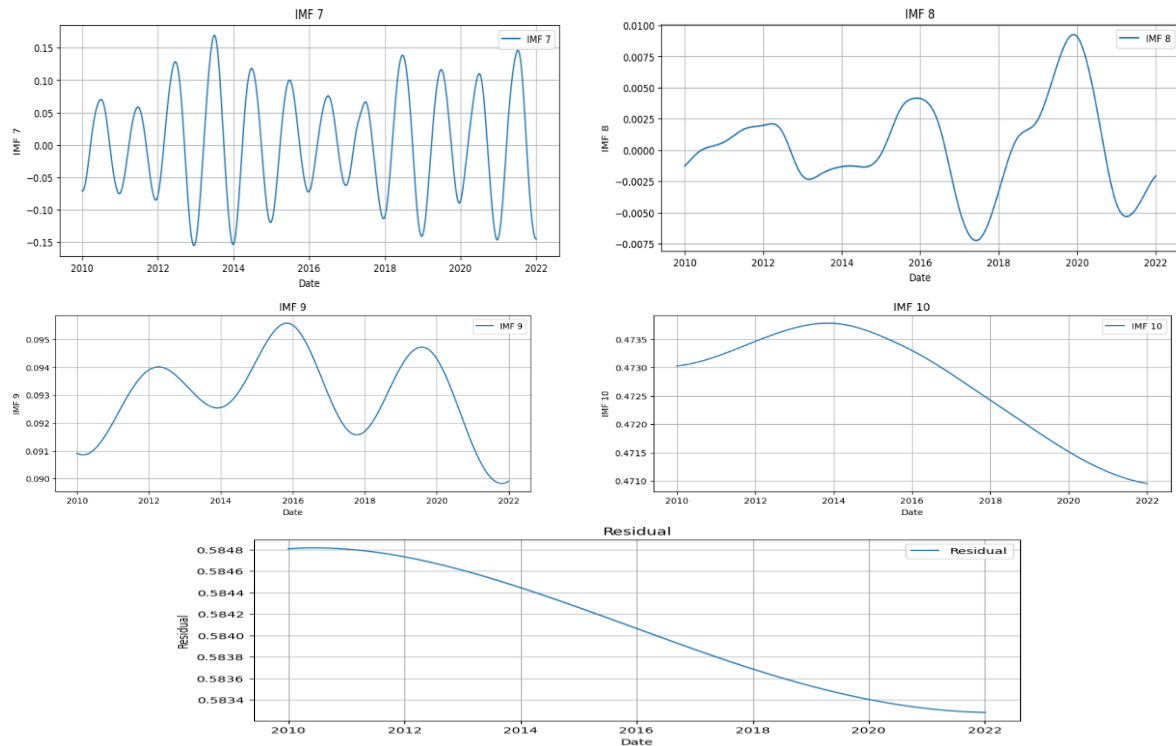


Figure 3 : Résultats de la décomposition EEMD dans la région de Biskra.

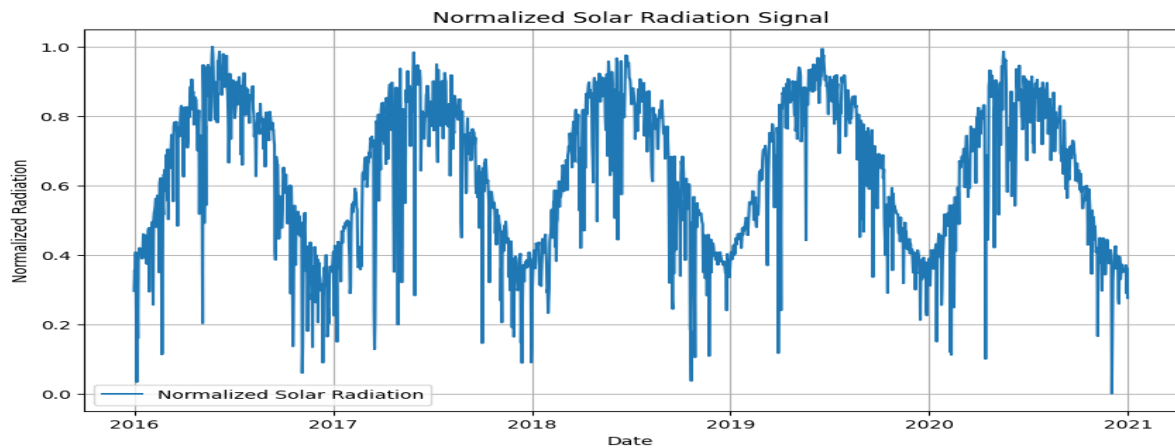
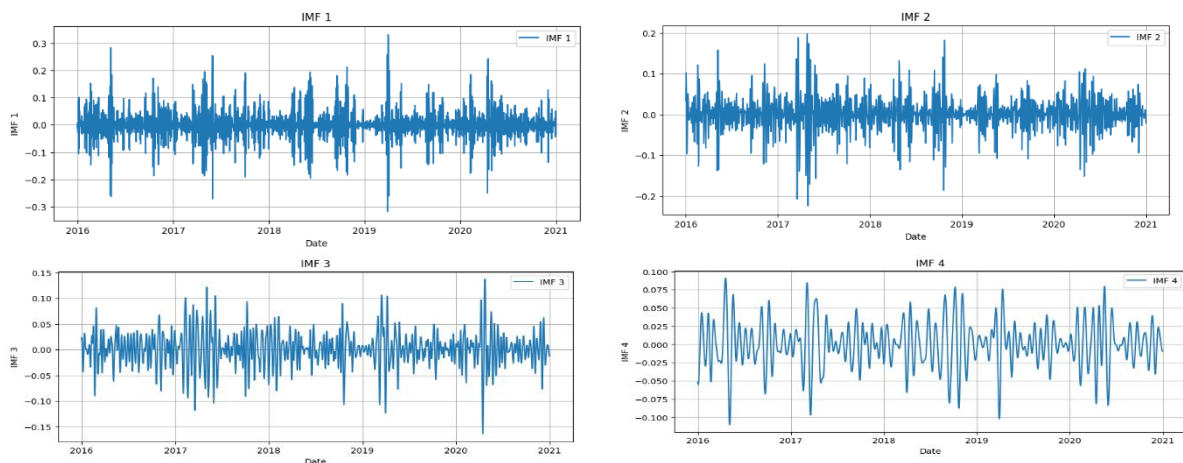


Figure 4 : Signal de Rayonnement Solaire Normalisé dans la région de Adrar.



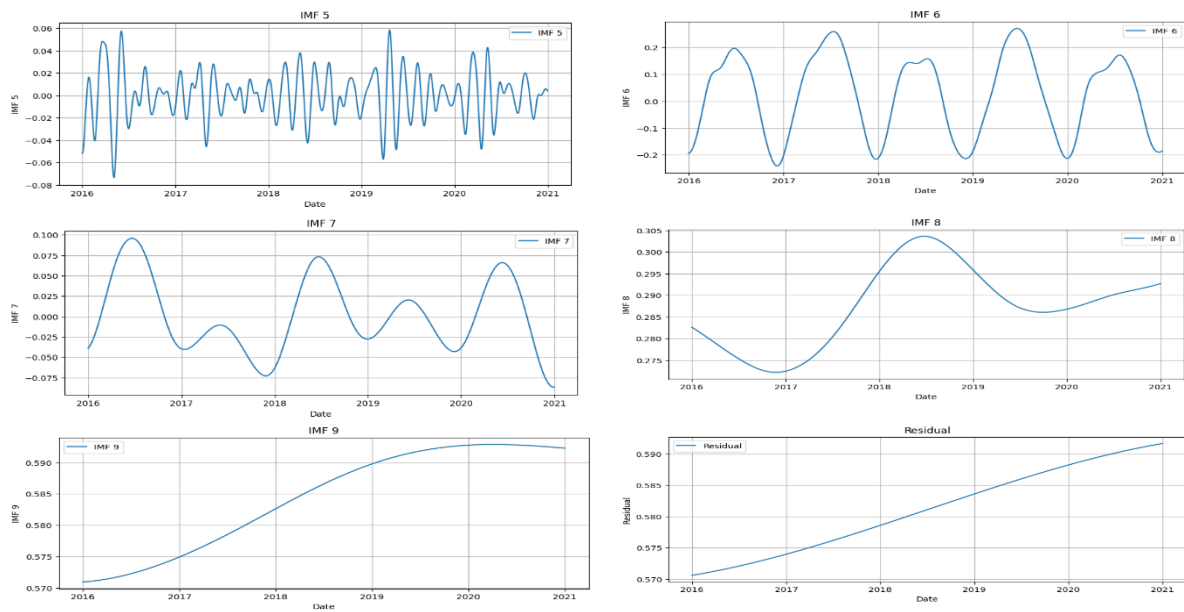


Figure 5 : Résultats de la décomposition EEMD dans la région d'Adrar.

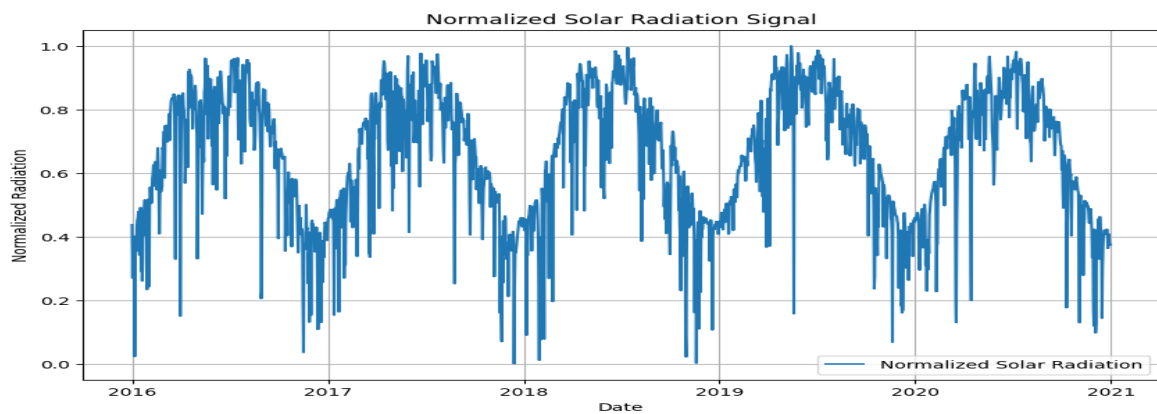
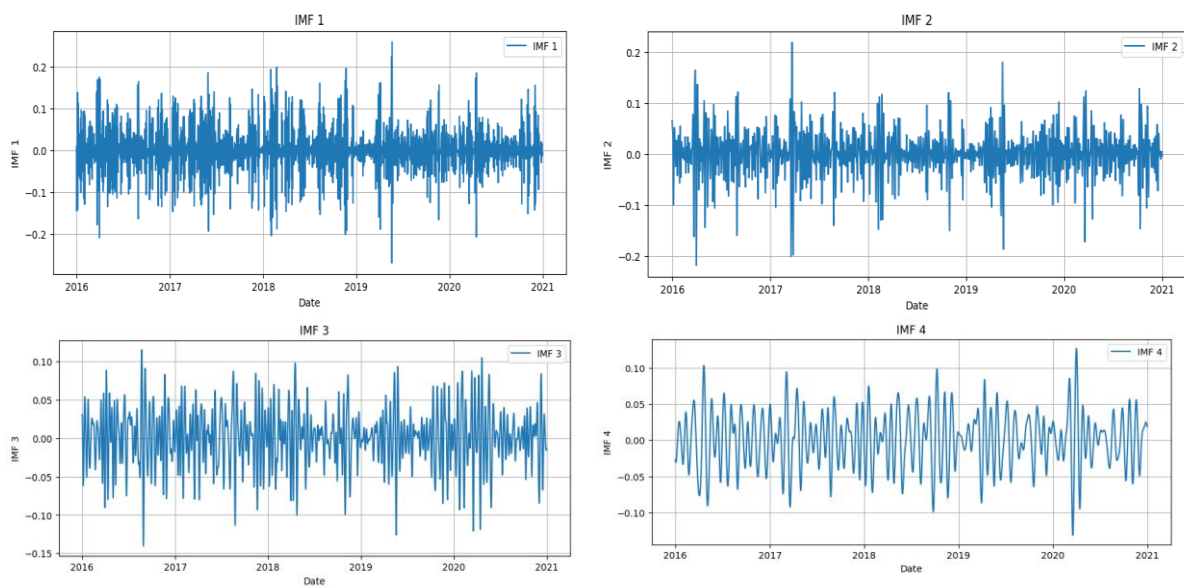


Figure 6 : Signal de Rayonnement Solaire Normalisé dans la région de Tamanrasset.



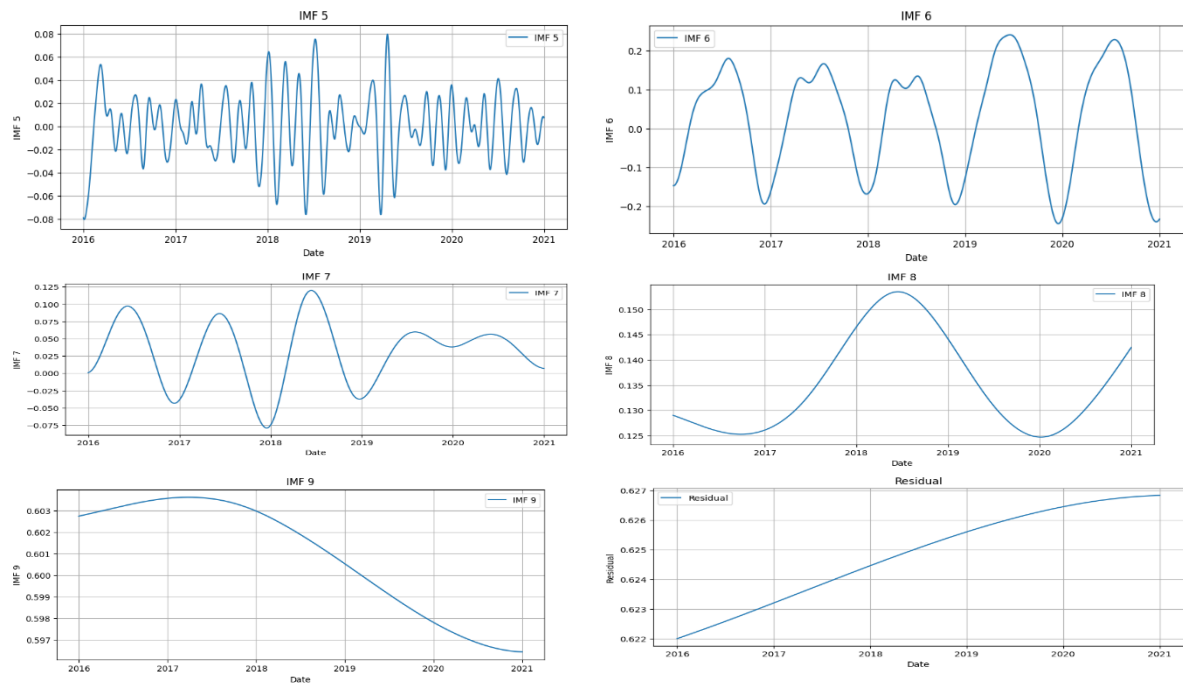


Figure 7 : Résultats de la décomposition EEMD dans la région de Tamanrasset.

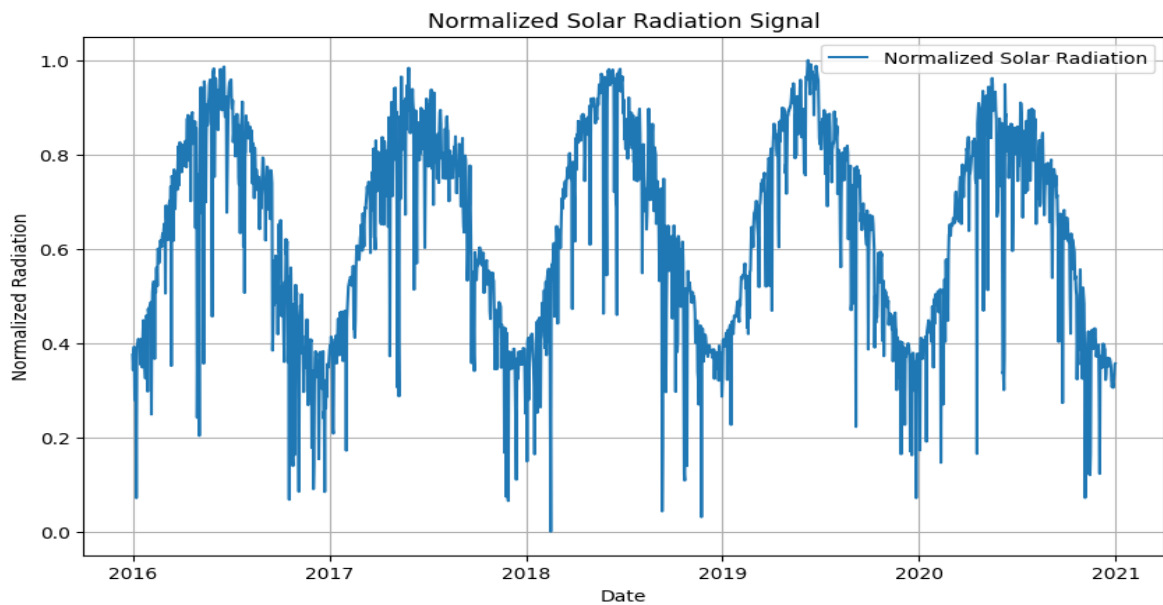
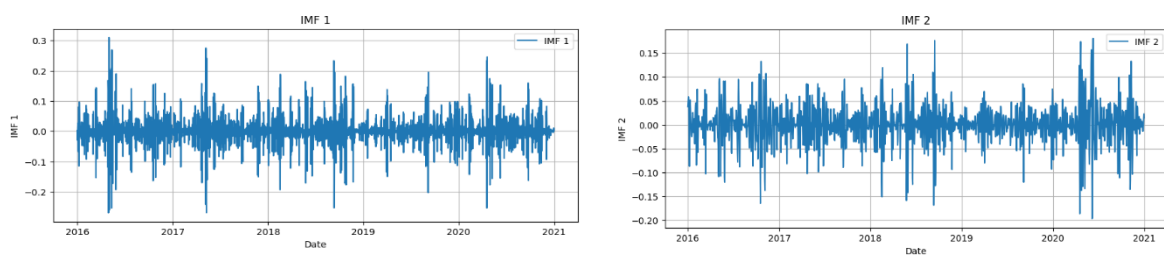


Figure 8 : Signal de Rayonnement Solaire Normalisé dans la région de Tindouf.



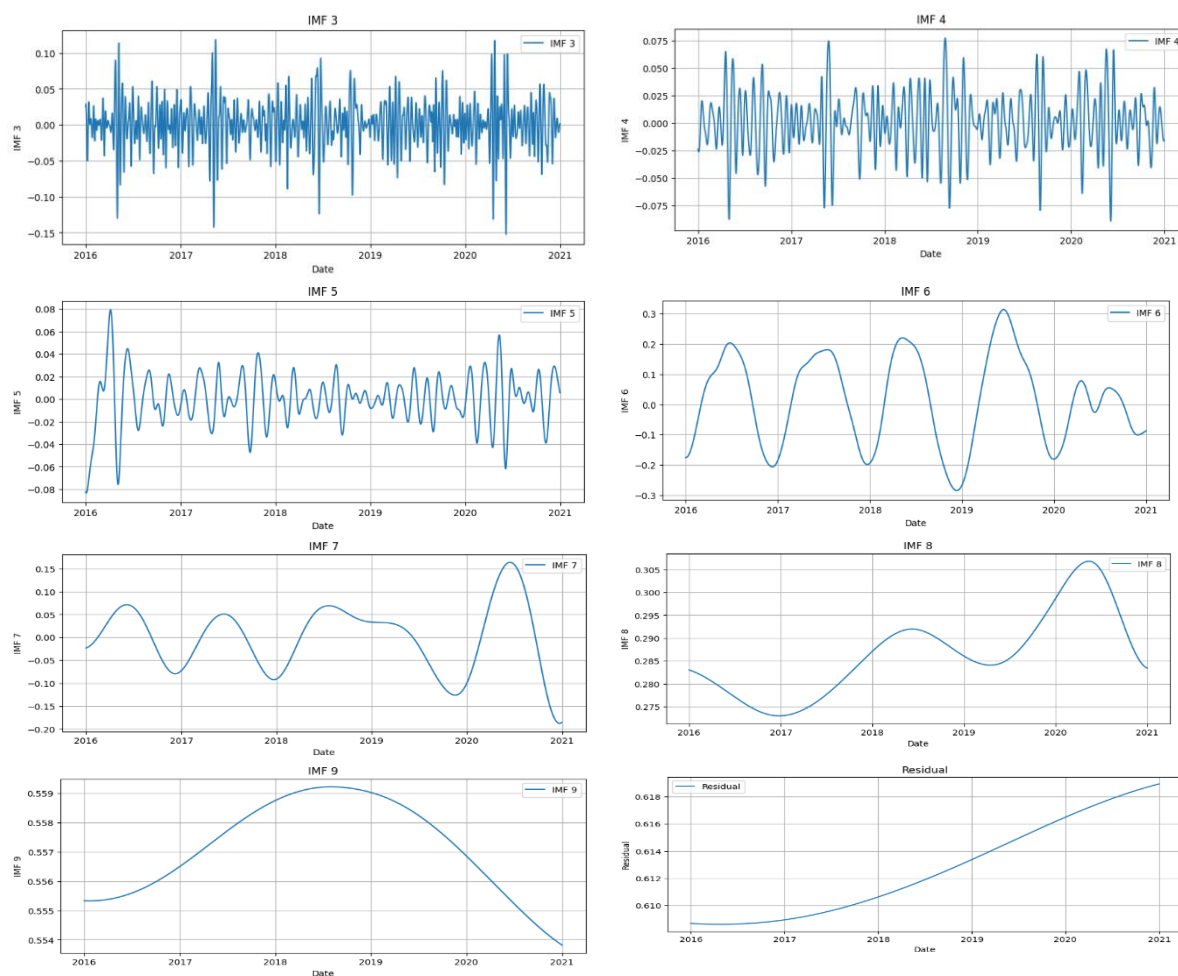
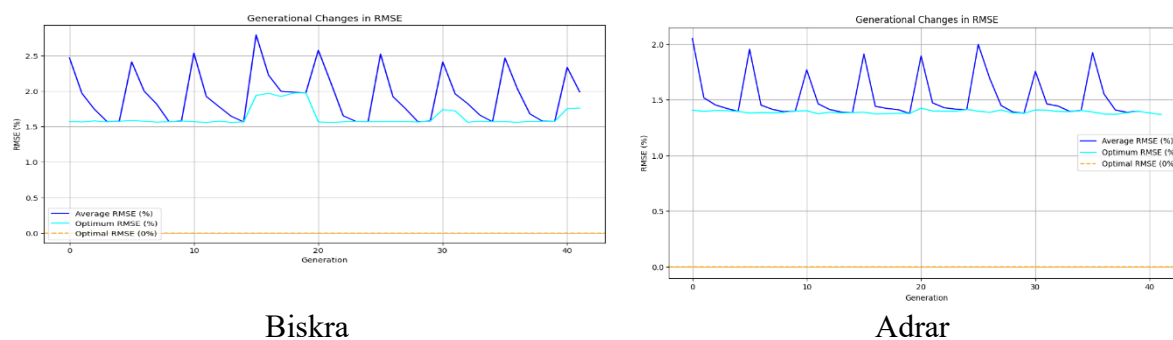


Figure 9: Résultats de la décomposition EEMD dans la région de Tindouf.

Les résultats de la décomposition EEMD fournissent une compréhension approfondie des dynamiques du rayonnement solaire dans les différentes régions étudiées. Cette analyse permet non seulement d'améliorer la précision des prédictions futures, mais également de mieux comprendre les mécanismes sous-jacents influençant les variations du rayonnement solaire.



Biskra

Adrar

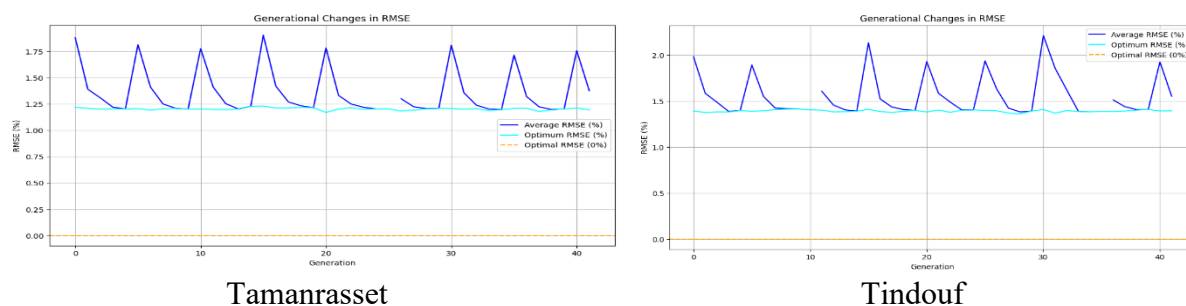
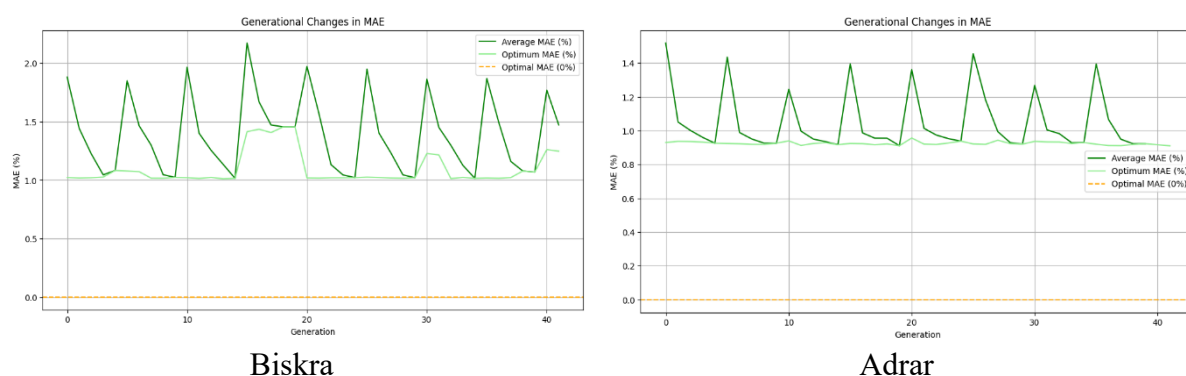


Figure 10 : Changement global du RMSE dans les régions de Biskra, Adrar, Tamanrasset et Tindouf.

La Figure 10 illustre le changement global du RMSE (Root Mean Squared Error) dans quatre régions spécifiques : Biskra, Adrar, Tamanrasset et Tindouf.

Biskra : Les résultats montrent un RMSE relativement faible, ce qui indique que les prédictions du modèle EEMD-GA-LSTM sont particulièrement précises. Cette précision peut être attribuée à une variabilité moindre du rayonnement solaire dans cette région, permettant ainsi au modèle d'ajuster plus efficacement les séries temporelles.

Adrar, Tamanrasset et Tindouf : En revanche, le RMSE dans ces régions est légèrement plus élevé par rapport à celui de Biskra. Cela suggère une légère diminution de la précision des prédictions, probablement en raison de conditions climatiques plus variables ou de fluctuations plus complexes du rayonnement solaire. Malgré cette augmentation du RMSE, il est important de noter que le modèle maintient néanmoins une bonne performance prédictive. Cela indique que, bien que les conditions dans ces régions soient plus difficiles à modéliser, la précision des prédictions reste globalement satisfaisante.



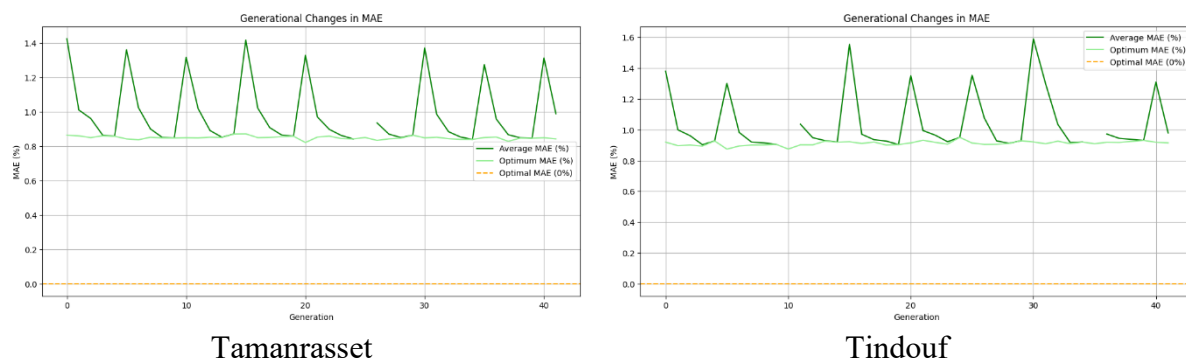
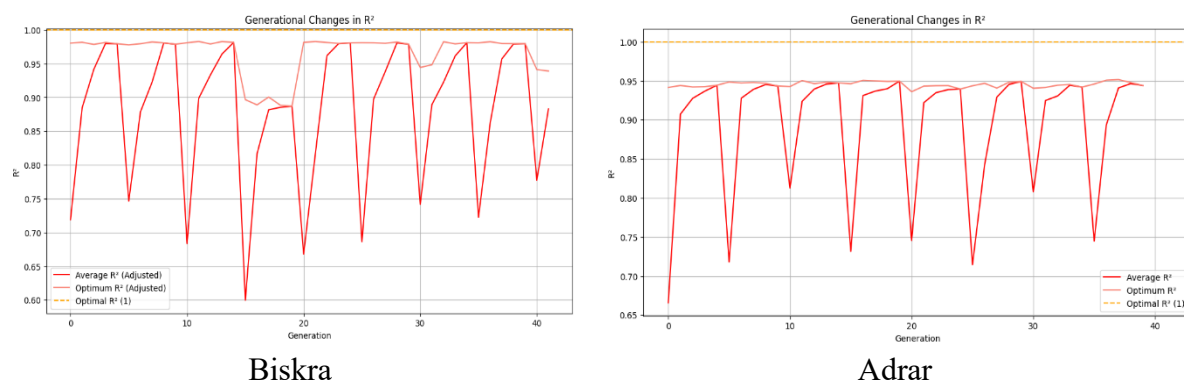


Figure 11 : Changement global du MAE dans les régions de Biskra, Adrar, Tamanrasset et Tindouf.

La Figure 11 illustre l'évolution du MAE (Mean Absolute Error) dans les régions de Biskra, Adrar, Tamanrasset et Tindouf. Le MAE représente la moyenne des erreurs absolues entre les valeurs prédites et observées, permettant d'évaluer l'écart moyen entre les prédictions du modèle et les données réelles. Un MAE faible est synonyme de prédictions plus précises.

Biskra : Dans cette région, le MAE est relativement bas, ce qui indique que le modèle EEMD-GA-LSTM génère des prédictions avec des erreurs moyennes limitées par rapport aux valeurs observées du rayonnement solaire. Cela témoigne de la capacité du modèle à capturer efficacement les variations du rayonnement dans un contexte climatique moins complexe.

Adrar, Tamanrasset et Tindouf : Le MAE est légèrement supérieur à celui de Biskra, suggérant une augmentation modeste de l'erreur moyenne dans les prédictions. Cette différence peut être liée à une variabilité climatique accrue ou à une complexité plus importante des séries temporelles dans ces régions. Néanmoins, l'augmentation du MAE reste raisonnable, montrant que le modèle maintient une performance satisfaisante malgré ces défis.



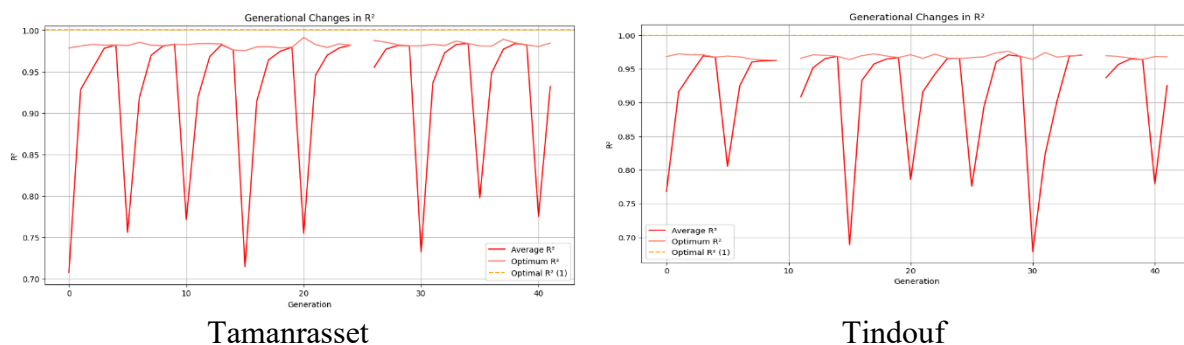
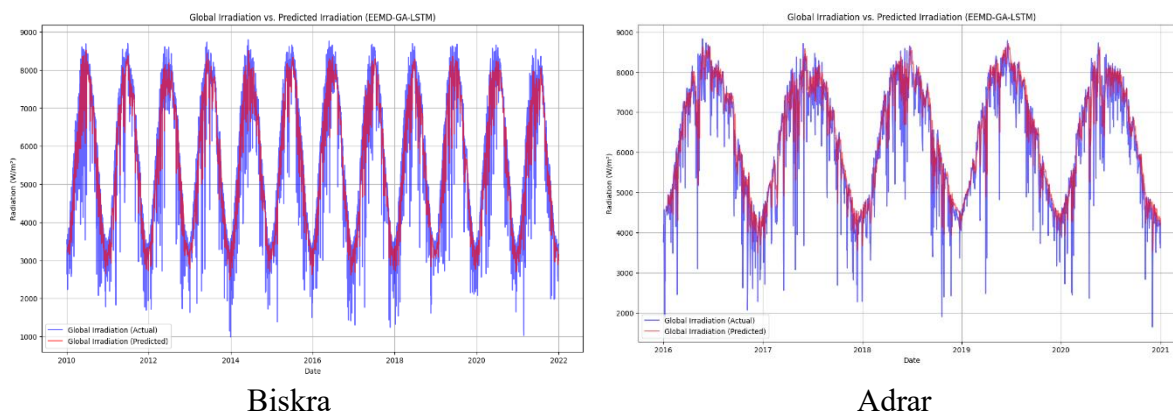


Figure 12 : Changement global du coefficient de détermination (R^2) dans les régions de Biskra, Adrar, Tamanrasset et Tindouf.

A Figure 12 illustre le changement global du coefficient de détermination (R^2) dans les régions de Biskra, Adrar, Tamanrasset et Tindouf. Le coefficient de détermination (R^2) est une mesure statistique indiquant la proportion de la variance des données observées qui est expliquée par le modèle. Plus R^2 est proche de 1, plus le modèle est efficace pour représenter les variations des données. En revanche, un R^2 proche de 0 suggère que le modèle explique peu de la variabilité observée.

Biskra : Dans cette région, le R^2 est élevé, se rapprochant de 1. Cela indique que le modèle EEMD-GA-LSTM réussit à expliquer presque toutes les variations du rayonnement solaire. Cette forte corrélation entre les valeurs prédites et observées suggère que le modèle est très précis à Biskra, où les fluctuations du rayonnement sont probablement plus régulières et moins influencées par des perturbations climatiques extrêmes.

Adrar, Tamanrasset et Tindouf : Bien que le R^2 pour ces régions soit légèrement inférieur à celui de Biskra, il reste élevé. Cela indique que le modèle continue d'expliquer efficacement la variance des données dans ces régions. Malgré une légère augmentation des erreurs dans ces régions, les performances du modèle demeurent satisfaisantes.



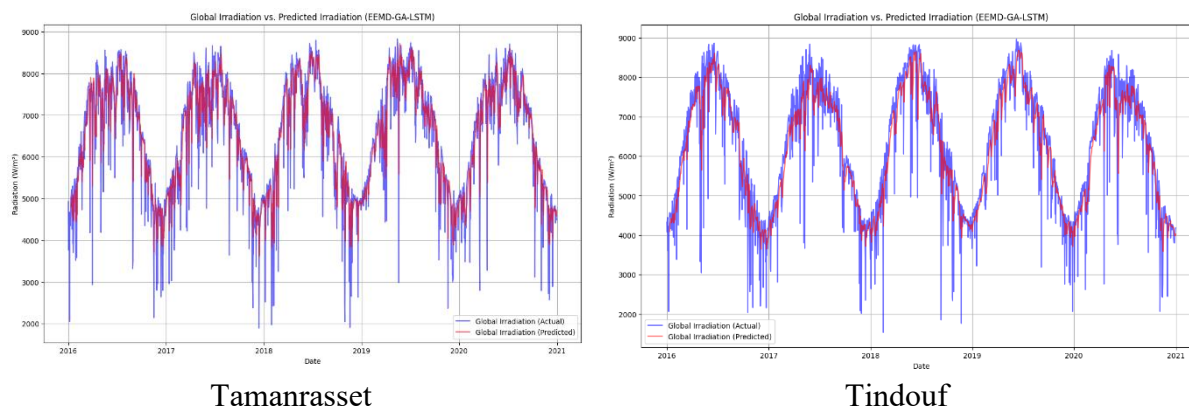


Figure 13 : Irradiation globale observée vs prédite en utilisant l'approche EEMD-GA-LSTM dans les régions de Biskra, Adrar, Tamanrasset et Tindouf

La Figure 13 présente une comparaison entre les valeurs observées et les valeurs prédites du rayonnement solaire obtenues par la méthode EEMD-GA-LSTM dans les quatre régions analysées : Biskra, Adrar, Tamanrasset, et Tindouf. Cette figure offre une évaluation visuelle de la précision des prédictions du modèle en opposant les valeurs prédites aux données réelles.

Biskra : Les valeurs prédites s'alignent étroitement avec les valeurs observées, ce qui témoigne d'une excellente précision du modèle dans cette région. Les courbes représentant les prédictions et les observations sont presque superposées, ce qui démontre que le modèle EEMD-GA-LSTM parvient à capturer efficacement les variations du rayonnement solaire. Ces résultats corroborent les performances antérieurement établies (RMSE, MAE, et R) qui soulignaient la robustesse du modèle à Biskra.

Adrar, Tamanrasset et Tindouf : Dans ces régions, les valeurs observées et prédites présentent également un bon alignement. Toutefois, des écarts légers peuvent être observés à certains moments, probablement en raison de la variabilité plus importante des conditions météorologiques. Malgré ces petites différences, la précision des prédictions demeure élevée, et le modèle continue de s'adapter efficacement aux données de ces régions.

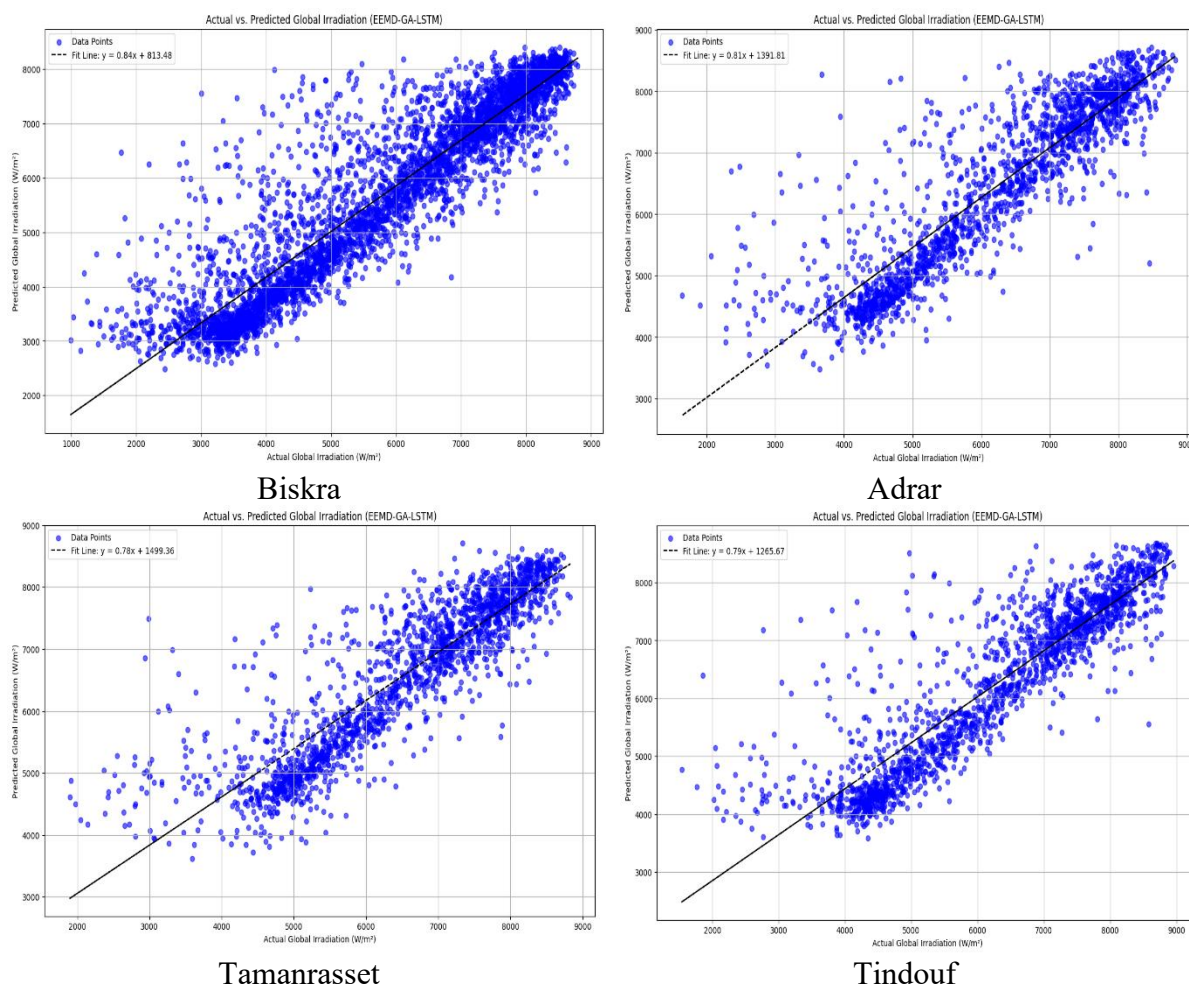


Figure 14 : Irradiation prédite en utilisant l'approche EEMD-GA-LSTM en fonction de l'irradiation actuelle dans les régions de Biskra, Adrar, Tamanrasset et Tindouf.

La Figure 14 illustre la relation entre les valeurs prédites et observées du rayonnement solaire, utilisant l'approche EEMD-GA-LSTM dans les régions de Biskra, Adrar, Tamanrasset, et Tindouf. Ce type de représentation visuelle permet d'évaluer la précision du modèle en observant l'alignement des prédictions avec les valeurs réelles. Idéalement, si le modèle était parfait, les points affichés sur le graphique devraient se situer le long de la ligne de régression avec une pente de 1, indiquant une correspondance parfaite entre les valeurs observées et prédites.

Biskra : Dans cette région, la majorité des points se rapprochent de la ligne de régression idéale, suggérant que le modèle EEMD-GA-LSTM prédit avec une grande précision les valeurs du rayonnement solaire. Les points sont bien alignés le long de la diagonale, ce qui indique que les écarts entre les valeurs réelles et prédites sont minimes. Cela renforce la performance du modèle dans cette région, comme observé dans les figures précédentes.

Adrar, Tamanrasset et Tindouf : Les points dans ces régions demeurent globalement alignés avec la ligne de régression, bien que quelques points présentent des écarts plus marqués, signalant des erreurs de prédiction dans certains cas. Néanmoins, la relation entre les valeurs observées et prédites reste très forte, ce qui démontre que le modèle continue de bien fonctionner, malgré une légère dégradation par rapport à Biskra.

Tableau1 : Résultat des indicateurs statistiques.

	<i>Biskra</i>	<i>Adrar</i>	<i>Tamanrasset</i>	<i>Tindouf</i>
<i>RMSE %</i>	<i>1.554</i>	<i>1.478</i>	<i>1.246</i>	<i>1.423</i>
<i>MAE %</i>	<i>1.074</i>	<i>0.985</i>	<i>0.828</i>	<i>0.786</i>
<i>R</i>	<i>0.986</i>	<i>0.964</i>	<i>0.976</i>	<i>0.972</i>

Le Tableau 1 présente les indicateurs statistiques RMSE (%), MAE (%) et R pour évaluer la performance du modèle EEMD-GA-LSTM dans les régions sahariennes de Biskra, Adrar, Tamanrasset et Tindouf. Les résultats montrent une bonne précision globale, avec des erreurs RMSE et MAE faibles et des coefficients de détermination R élevés. Tamanrasset affiche l'erreur RMSE la plus faible (1.246 %), indiquant une prédiction plus précise dans cette région, tandis que Biskra présente la plus grande erreur (1.554 %), bien que restant relativement faible. Les erreurs MAE suivent la même tendance, avec Tindouf et Tamanrasset affichant les valeurs les plus basses (0.786 % et 0.828 % respectivement), ce qui suggère une précision plus élevée du modèle dans ces zones. En ce qui concerne le coefficient de détermination R, Biskra atteint la meilleure valeur (0.986), indiquant un très bon ajustement du modèle, tandis que Adrar présente la valeur la plus faible (0.964), bien qu'elle reste excellente. Ces résultats confirment la robustesse du modèle EEMD-GA-LSTM, capable de fournir des prédictions fiables du rayonnement solaire dans les régions sahariennes, même avec des bases de données limitées.

Tableau 2 : Comparaison des performances des modèles présentés avec les résultats de diverses études.

<i>Référence</i>	<i>modèle</i>	<i>Région</i>	<i>R</i>
<i>[12]</i>	<i>ANN-GA</i>	<i>Adrar</i>	<i>0.9962</i>
<i>[12]</i>	<i>ANN-PSO</i>	<i>Adrar</i>	<i>0.9929</i>
<i>[12]</i>	<i>ANFIS-GA</i>	<i>Adrar</i>	<i>0.9210</i>
<i>[12]</i>	<i>ANFIS-PSO</i>	<i>Adrar</i>	<i>0.9523</i>
<i>[12]</i>	<i>ANN-GA</i>	<i>Tamanrasset</i>	<i>0.9887</i>
<i>[12]</i>	<i>ANN-PSO</i>	<i>Tamanrasset</i>	<i>0.9987</i>
<i>[12]</i>	<i>ANFIS-GA</i>	<i>Tamanrasset</i>	<i>0.8425</i>
<i>[12]</i>	<i>ANFIS-PSO</i>	<i>Tamanrasset</i>	<i>0.9877</i>
<i>[12]</i>	<i>ANN-GA</i>	<i>Tindouf</i>	<i>0.9998</i>
<i>[12]</i>	<i>ANN-PSO</i>	<i>Tindouf</i>	<i>0.9999</i>

[12]	ANFIS-GA	Tindouf	0.9866
[12]	ANFIS-PSO	Tindouf	0.9986
Cette étude	EEMD-GA-LSTM	Biskra	0.986
Cette étude	EEMD-GA-LSTM	Adrar	0.964
Cette étude	EEMD-GA-LSTM	Tamanrasset	0.976
Cette étude	EEMD-GA-LSTM	Tindouf	0.972

Le tableau compare la performance du modèle EEMD-GA-LSTM, développé dans cette étude, avec plusieurs modèles issus d'une étude antérieure [12], notamment ANN-GA, ANN-PSO, ANFIS-GA et ANFIS-PSO, sur différentes régions sahariennes (Adrar, Tamanrasset et Tindouf). Les résultats montrent que les modèles ANN-GA et ANN-PSO atteignent des valeurs de R très élevées, supérieures à 0.98, indiquant une très bonne précision. En revanche, les modèles ANFIS présentent plus de variabilité, en particulier ANFIS-GA à Tamanrasset, où R chute à 0.8425, suggérant une performance moins fiable dans cette région. À Tindouf, tous les modèles atteignent des R proches de 1, indiquant une meilleure capacité de prévision dans cette zone.

Concernant le modèle EEMD-GA-LSTM, l'intervalle de R obtenu (0.964 - 0.986) montre une performance compétitive par rapport aux autres approches, avec une stabilité accrue sur plusieurs régions. Bien que ses valeurs soient légèrement inférieures aux meilleures performances de ANN-GA ou ANN-PSO, ce modèle présente l'avantage de fonctionner efficacement même avec un ensemble de données limité, ce qui est crucial dans les régions sahariennes où les bases de données météorologiques sont souvent réduites. Ainsi, bien que les méthodes ANN et ANFIS optimisées par GA et PSO offrent parfois de meilleures performances ponctuelles, EEMD-GA-LSTM constitue une approche plus équilibrée, robuste et adaptée aux conditions sahariennes.

Conclusion

En conclusion, le modèle EEMD-GA-LSTM proposé s'est révélé être une approche performante pour la prédiction du rayonnement solaire dans les quatre régions étudiées de Biskra, Tamanrasset, Adrar et Tindouf. Les résultats obtenus, évalués à l'aide de l'indice de performance R, attestent d'une précision et d'une fiabilité élevées, comparables à celles d'autres méthodes existantes en matière de prévision et de précision. Un aspect particulièrement notable est la robustesse du modèle, qui a démontré une capacité à fournir des prédictions fiables même avec un volume limité de données. Cette caractéristique est cruciale pour son application dans

des régions où les infrastructures météorologiques sont peu développées ou où les données sont rares.

Le modèle a affiché des performances quasi parfaites pour la région de Biskra, avec une excellente corrélation entre les valeurs observées et prédites. Pour les régions d'Adrar, Tamanrasset et Tindouf, des écarts mineurs ont été observés, principalement attribuables à la plus grande variabilité des conditions météorologiques locales. Néanmoins, le modèle a maintenu un niveau de précision très élevé, confirmant sa capacité à s'adapter à des environnements complexes.

Ces résultats positionnent le modèle EEMD-GA-LSTM comme un outil particulièrement adapté aux zones disposant de peu de données météorologiques, tout en offrant des estimations fiables du rayonnement solaire. Cette capacité est essentielle pour des applications dans des régions reculées ou isolées, où une planification énergétique précise est indispensable. En somme, ce chapitre confirme le potentiel du modèle en tant qu'outil innovant et efficace pour l'estimation du rayonnement solaire, ouvrant des perspectives prometteuses pour son utilisation dans des contextes variés, notamment dans les régions où l'accès à des données détaillées reste limité.

Références

- [1] Ouedraogo, S. (2023). Développement de Stratégies Optimisées de Gestion de l'Energie Intermittente dans un Micro Réseau Photovoltaïque avec Stockage (Doctoral dissertation, Université Pascal Paoli).
- [2] Chen, Y., Dong, Z., Wang, Y., Su, J., Han, Z., Zhou, D., ... & Bao, Y. (2021). Short-term wind speed predicting framework based on EEMD-GA-LSTM method under large scaled wind history. *Energy Conversion and Management*, 227, 113559.
- [3] Awajan, A. M., Ismail, M. T., & Wadi, S. A. L. (2019). A review on empirical mode decomposition in forecasting time series. *Italian Journal of Pure and Applied Mathematics*, 43, 301-323.
- [4] Bakhtiari, S. (2012). New empirical mode decompositions and their applications in signal and image processing. The University of Texas at San Antonio.
- [5] Xu, P., Aamir, M., Shabri, A., Ishaq, M., Aslam, A., & Li, L. (2020). A new approach for reconstruction of IMFs of decomposition and ensemble model for forecasting crude oil prices. *Mathematical Problems in Engineering*, 2020(1), 1325071.
- [6] Bouakline, O., Elidrissi, A., Khomsi, K., & Leghrib, R. (2024). A hybrid deep learning model-based LSTM and modified genetic algorithm for air quality applications. *Environmental Monitoring and Assessment*, 196(12), 1-25.

-
- [7] Okasha, N. M., & Frangopol, D. M. (2009). Lifetime-oriented multi-objective optimization of structural maintenance considering system reliability, redundancy and life-cycle cost using GA. *Structural Safety*, 31(6), 460-474.
- [8] Silva, C. A., Sousa, J. M. C., & Runkler, T. A. (2008). Rescheduling and optimization of logistic processes using GA and ACO. *Engineering Applications of Artificial Intelligence*, 21(3), 343-352.
- [9] Ding, L., Yue*, Y., Ahmet, K., Jackson, M., & Parkin, R. (2005). Global optimization of a feature-based process sequence using GA and ANN techniques. *International Journal of Production Research*, 43(15), 3247-3272.
- [10] Kumari, P., & Toshniwal, D. (2021). Long short term memory–convolutional neural network based deep hybrid approach for solar irradiance forecasting. *Applied Energy*, 295, 117061.
- [11] Gupta, A., & Gupta, K. (2022). Short term solar irradiation prediction framework based on EEMD-GA-LSTM Method. *Strategic Planning for Energy and the Environment*, 255-280.
- [12] Senouci, A., Benatallah, A., Benatallah, D., Bouchouicha, K., & Maouedj, R. (2025). EEMD-GA-LSTM: An innovative approach for daily solar radiation estimation in southern of Algeria. *Edelweiss Applied Science and Technology*, 9(1), 278-291.
- [13] Ullah, K., Ahsan, M., Hasanat, S. M., Haris, M., Yousaf, H., Raza, S. F., ... & Ullah, Z. (2024). Short-term load forecasting: A comprehensive review and simulation study with CNN-LSTM hybrids approach. *IEEE Access*.
- [14] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, 31(7), 1235-1270.
- [15] Ma, Z., & Cai, H. (2024, May). Research on Electricity Load Forecasting Based on EEMD-GA-LSTM. In *2024 36th Chinese Control and Decision Conference (CCDC)* (pp. 4434-4439). IEEE.
- [16] Zeng, Z., Yang, H., Zhao, R., & Meng, J. (2013). Nonlinear characteristics of observed solar radiation data. *Solar Energy*, 87, 204-218. [18] Selahdja, B. (2024). The Geographical distribution of prehistoric sites in the region of Biskra. *Journal El-Baheth in Human and Social Sciences Volume15 Number*, 135-147.
- [17] Cheng, C., Sa-Ngasoongsong, A., Beyca, O., Le, T., Yang, H., Kong, Z., & Bukkapatnam, S. T. (2015). Time series forecasting for nonlinear and non-stationary processes: a review and comparative study. *Iie Transactions*, 47(10), 1053-1071.
- [18] Lei, Y., He, Z., & Zi, Y. (2009). Application of the EEMD method to rotor fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, 23(4), 1327-1338.
- [19] Jiang, H., Li, C., & Li, H. (2013). An improved EEMD with multiwavelet packet for rotating machinery multi-fault diagnosis. *Mechanical Systems and Signal Processing*, 36(2), 225-239.

- [20] Osman, H., Ghafari, M., & Nierstrasz, O. (2017, February). Hyperparameter optimization to improve bug prediction accuracy. In 2017 IEEE workshop on machine learning techniques for software quality evaluation (MaLTaSQuE) (pp. 33-38). IEEE.
- [21] Asiri, A. A., Shaf, A., Ali, T., Aamir, M., Irfan, M., & Alqahtani, S. (2024). Enhancing brain tumor diagnosis: an optimized CNN hyperparameter model for improved accuracy and reliability. *PeerJ Computer Science*, 10, e1878.
- [22] Diaz, G. I., Fokoue-Nkoutche, A., Nannicini, G., & Samulowitz, H. (2017). An effective algorithm for hyperparameter optimization of neural networks. *IBM Journal of Research and Development*, 61(4/5), 9-1.
- [23] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- [24] Mellit, A., & Pavan, A. M. (2010). A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Solar energy*, 84(5), 807-821.
- [25] Wang, W. C., Chau, K. W., Xu, D. M., & Chen, X. Y. (2015). Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. *Water resources management*, 29, 2655-2675.
- [26] Al Samrout, M. (2024). *Approches mono et bi-objective pour l'optimisation intégrée des postes d'amarrage et des grues de quai dans les opérations de transbordement* (Doctoral dissertation, Normandie).
- [27] Khessiba, S. (2024). *Stratégies d'optimisation des hyper-paramètres de réseaux de neurones appliqués aux signaux temporels biomédicaux* (Doctoral dissertation, Institut Polytechnique de Paris; Université de Sousse (Tunisie)).
- [28] Kumar, M., Husain, D. M., Upreti, N., & Gupta, D. (2010). Genetic algorithm: Review and application. *Available at SSRN 3529843*.
- [29] Mathew, T. V. (2012). Genetic algorithm. *Report submitted at IIT Bombay*, 53, 18-19.
- [30] Mirjalili, S., & Mirjalili, S. (2019). Genetic algorithm. *Evolutionary algorithms and neural networks: Theory and applications*, 43-55.
- [31] Lambora, A., Gupta, K., & Chopra, K. (2019, February). Genetic algorithm-A literature review. In 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon) (pp. 380-384). IEEE.



Conclusion Générale



Conclusion générale

Dans un contexte mondial marqué par la transition énergétique et la recherche de sources d'énergie renouvelable fiables, l'exploitation du rayonnement solaire s'impose comme une solution stratégique, particulièrement dans les régions sahariennes caractérisées par une insolation exceptionnelle. Toutefois, l'estimation précise du rayonnement solaire dans ces zones extrêmes représente un défi majeur, en raison de la variabilité spatio-temporelle, du manque de données fiables, et des conditions climatiques spécifiques comme les tempêtes de sable.

Dans cette optique, la présente thèse a été consacrée à l'étude, la modélisation et l'optimisation de méthodes avancées d'estimation du rayonnement solaire global dans les régions sahariennes d'Algérie, à travers une approche méthodologique rigoureuse combinant les bases théoriques, les outils informatiques et l'intelligence artificielle appuyée par des métaheuristiques.

Dans un premier temps, une analyse approfondie des paramètres influençant le rayonnement solaire (facteurs géographiques, astronomiques et climatiques) a été réalisée, mettant en lumière l'importance de la qualité des données et des méthodes d'estimation. L'étude des méthodes de mesure traditionnelles et modernes a démontré l'évolution continue des approches d'acquisition de données solaires, soulignant la nécessité de recourir à des modèles robustes capables de compenser les incertitudes.

Ensuite, une démarche hybride basée sur l'utilisation des modèles ANN-GA, ANN-PSO, ANFIS-GA et ANFIS-PSO a été développée et testée. L'analyse des résultats a montré que les modèles ANN-GA et ANN-PSO ont permis d'obtenir des performances remarquables pour la prédiction du rayonnement solaire :

- Le modèle ANN-GA s'est avéré être le plus performant pour la région d'Adrar, atteignant un MPE extrêmement faible (0,0015 %), un RMSE bas (3,2554 %) et un coefficient de corrélation élevé ($R = 0,9962$).
- Le modèle ANN-PSO, quant à lui, a surpassé les autres pour les régions de Tindouf et Tamanrasset, avec des performances exceptionnelles ($R = 0,9999$ pour Tindouf et $R = 0,9987$ pour Tamanrasset), confirmant ainsi la pertinence de l'optimisation par essaim de particules dans la calibration des réseaux de neurones.

Parallèlement, une méthode innovante, le modèle EEMD-GA-LSTM, a été mise en œuvre pour intégrer la décomposition du signal (EEMD), l'optimisation par algorithme génétique (GA) et la capacité prédictive des réseaux neuronaux à mémoire longue (LSTM). Les résultats obtenus dans les régions de Biskra, Adrar, Tamanrasset et Tindouf ont démontré :

- Une bonne précision générale, avec des erreurs RMSE et MAE faibles (par exemple, RMSE de 1,246 % à Tamanrasset) ;
- Des coefficients de détermination R élevés, variant entre 0,964 et 0,986, ce qui indique une excellente capacité de modélisation même dans des environnements à bases de données limitées.

La comparaison des modèles a révélé que, si les approches ANN-GA et ANN-PSO offrent localement de meilleures performances, notamment en termes de R très proches de 1, le modèle EEMD-GA-LSTM présente quant à lui l'avantage d'une stabilité accrue sur plusieurs sites et d'une robustesse face aux jeux de données plus restreints, ce qui est crucial dans le contexte saharien.

À l'avenir, nous explorerons de nouvelles méthodes hybrides afin d'améliorer les modèles de prévision du rayonnement solaire tout en garantissant leur précision et leur flexibilité. En poursuivant cet objectif, nous espérons contribuer à un avenir plus durable et respectueux de l'environnement en fournissant aux parties concernées des informations et des ressources essentielles pour une gestion optimale de l'énergie solaire.