

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

**UNIVERSITE DE RELIZANE**

Faculté des Sciences et Technologies  
Département de Génie Mécanique



**THESE DE DOCTORAT LMD 3<sup>ème</sup> cycle**

Filière : Génie Mécanique

Spécialité : Énergétique

Présentée par

**DAHMANI Abdennasser**

Titre de thèse

**Contribution à la modélisation de l'irradiation solaire par  
l'intelligence artificielle. Etude comparative.**

Soutenue le : **06/01/2024**

**Devant le jury composé de :**

Président	HADJ MOSTEFA Adda	MCA	Univ.de Relizane
Directeur de thèse	AMMI Yamina	MCA	Univ.de Médéa
Co-Directeur de thèse	HANINI Salah	Pr	Univ.de Médéa
Examineur	BOUDAHRI Fethi	Pr	Univ.de Tlemcen
Examineur	BABA AHMED Abderrazzak	MCA	Univ.de Relizane
Invité	LAIDI Mamaar	Pr	Univ.de Médéa

Année Universitaire : 2023/2024

## *Dédicaces*

*Je dédie cette thèse à :*

*Mes très chers et aimables parents.*

*A ma famille et mes collègues,*

*À tous ceux qui m'ont soutenu de près ou de loin.*

*Abdennasser DAHMANI*

## **Remerciements**

*Je tiens à remercier tout d'abord **ALLAH**, le tout puissant, de m'avoir donné la volonté et le courage pour réaliser ce travail.*

*Je tiens à exprimer ma gratitude envers ma directrice de thèse, Mademoiselle **AMMI Yamina**, Maître de Conférences (A) à l'Université de Médéa, pour son encadrement et sa supervision de cette recherche. Bien que son expertise ait été inestimable dans la formulation des questions et des objectifs de recherche, elle a également soutenu mes idées depuis le début et a cru en mes décisions.*

*Je souhaite exprimer ma plus profonde gratitude envers Monsieur le Professeur **HANINI Salah** de l'Université de Médéa. Il a été bien plus qu'un co-directeur de thèse, il a été un soutien inestimable, faisant preuve d'une implication totale dans tous les aspects de ce travail au cours des trois années que nous avons passées à travailler ensemble.*

*Je profite de l'occasion de la présentation de ce travail pour exprimer mes vifs remerciements à Monsieur **HADJ MOSTEFA Adda**, Maître Conférences (A) à l'Université de Relizane, pour avoir accepté de présider le jury de soutenance.*

*Je tiens également à présenter ma profonde gratitude à Monsieur **BOUDAHRI Fethi**, Professeur à l'Université de Tlemcen, Monsieur le Professeur **LAIIDI Mammaar** à l'Université de Médéa et Monsieur **BABA AHMED Abderrazzak**, Maître Conférences (A) à l'Université de Relizane, d'avoir accepté de juger ce modeste travail.*

*Je remercie vivement Monsieur **Nadjem Bailek**, Monsieur **Billel Amiri** et Mademoiselle **KOUIDRI Ikram** pour leur collaboration technique efficace et leur constante disponibilité.*

*Je tiens à exprimer ma profonde reconnaissance à Monsieur **Redha Mohamed Yaïche**, chercheur au CDER pour son aide précieuse.*

*Enfin, mes plus profonds remerciements vont à ma mère et à mon père. Merci pour l'écoute, le soutien et toutes vos marques d'attention. Je vous dois énormément et je ne serais certainement pas arrivé là aujourd'hui sans vous.*

## الملخص

### المساهمة في نمذجة الإشعاع الشمسي باستخدام الذكاء الاصطناعي. دراسة مقارنة

الهدف الذي تم تقديمه في هذا العمل هو التنبؤ بالإشعاع الشمسي الساعي الإجمالي الذي يتم استقباله على سطح أفقي باستخدام خوارزميات الذكاء الاصطناعي (AI) وتحديد آليات دعم القوى الناقلة (MVS) والشبكات العصبونية الاصطناعية (RNA) لأداء هذه المهمة.

يتكون هذا العمل من مساهمتين متميزتين. في المساهمة الأولى، تم تطوير منهجية لتحسين الشبكة العصبونية الاصطناعية من نوع الإدراك الحسي متعدد الطبقات (RNA-PMC) وتم تنفيذ آلية دعم القوى الناقلة مع وظيفة النواة من نوع القاعدة الشعاعية لنمذجة الإشعاع الشمسي الساعي باستخدام البيانات التي تضمنت العديد من المعلمات الجوية والتي تم جمعها من مدينتين في الجزائر (بوزريعة، تمنراست) وتم اختيارهما لتغطية مناخات مختلفة. أظهرت النتائج المحصلة أن النموذج MVS-RBF أفضل من النموذج

RNA-PMC بخوارزمية التعلم Broyden-Fletcher-Goldfarb-Shanno (BFGS) في تنبؤ الإشعاع الشمسي الأفقي الإجمالي مع قيمة معامل الارتباط  $R=0.9876$  وجذر الخطأ المربع المتوسط  $REQM=38.70 \text{ Wh/m}^2$  للمرحلة الكلية.

المساهمة الثانية تتمثل في تطوير نموذج الشبكات العصبونية المجتمعة (BRNA) صالحة لتنبؤ الإشعاع الشمسي الساعي في مدينة بوزريعة. تم توفير قاعدة البيانات التي تحتوي على العديد من المعلمات الجوية المقاسة من قبل مركز تطوير الطاقة المتجددة في بوزريعة-الجزائر. أظهرت النتائج المحصلة نتائج مرضية للغاية ودقيقة بمعامل ارتباط يبلغ  $R = 0.9680$  وجذر متوسط الخطأ المربعي  $REQM = 62.4856 \text{ Wh/m}^2$  وخطأ معياري في التنبؤ  $ESP = 12.79 \%$ . تشير النتائج المحصلة إلى أن نموذج الشبكات العصبونية المجتمعة (BRNA) أكثر تلائماً ومتانةً للتنبؤ الإشعاع الشمسي الساعي الإجمالي المستلم على سطح أفقي.

**كلمات مفتاحية:** ذكاء إصطناعي؛ نمذجة؛ الإشعاع الشمسي؛ دراسة مقارنة

## Summary

### **Contribution to solar irradiation modeling using artificial intelligence. A comparative study.**

The objective presented in this work is to predict hourly global solar irradiation (HGSR) received on a horizontal surface using artificial intelligence algorithms, specifically support vector machines (SVM) and artificial neural networks (ANN) to realize this task.

This work is composed of two distinct contributions. In the first contribution, a methodology was developed With a view to optimize the artificial neural network of type the multilayer perceptron type (ANN-MLP), and a support vector machine with a radial basis function kernel (SVM-RBF) was used to model hourly solar irradiation using data including multiple meteorological parameters collected from two cities in Algeria (Bouzaréah, Tamanrasset) and chosen to cover different climates. The results obtained showed that the SVM-RBF model better than the RNN-MLP model with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) learning algorithm for predicting global horizontal irradiation with a correlation coefficient  $R$  of 0.99 and a root mean square error (RMSE) of  $38.70 \text{ Wh/m}^2$  for the total phase.

The second contribution Consists of developing a bootstrap aggregated neural network (BANN) valid for predicting hourly solar irradiation in the city of Bouzaréah. The database containing various measured meteorological parameters was provided by the Center for Renewable Energy Development (CDER) in Bouzaréah, Algeria. The results obtained are highly satisfactory and accurate with a correlation coefficient of 96.80%, a root mean square error (RMSE) of  $62.4856 \text{ Wh/m}^2$ , and a standard prediction error (SPE) of 12.79%. The results obtained indicate that the BANN model is better suited and more robust for predicting hourly global solar irradiation received on a horizontal surface.

**Keywords:** Artificial Intelligence, Modeling, Solar Irradiation, Comparative Study.

## Résumé

### Contribution à la modélisation de l'irradiation solaire par l'intelligence artificielle. Etude comparative

L'objectif présenté à ce travail est de prédire l'irradiation solaire globale horaire (ISGH) reçue sur une surface horizontale en utilisant les algorithmes d'intelligence artificielle et plus précisément des machines à vecteur de support (MVS) et réseaux de neurones artificiels (RNA) pour réaliser cette tâche.

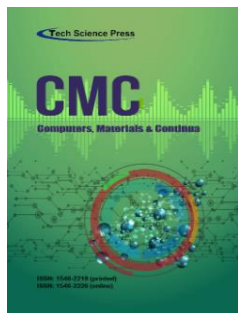
Ce travail est composé de deux contributions distinctes. Dans la première contribution, une méthodologie a été développée dans le but d'optimiser le réseau neuronal artificiel de type perceptron multicouche (ANN-MLP), et une machine à vecteurs de support avec un noyau de fonction de base radiale (SVM-RBF) a été utilisée pour modéliser l'irradiation solaire horaire en utilisant des données comprenant plusieurs paramètres météorologiques collectés dans deux villes en Algérie (Bouzaréah, Tamanrasset) choisies pour couvrir différents climats. Les résultats obtenus ont montré que le modèle SVM-RBF était meilleur que le modèle ANN-MLP avec l'algorithme d'apprentissage Broyden-Fletcher-Goldfarb-Shanno (BFGS) pour prédire l'irradiation horizontale globale avec un coefficient de corrélation  $R$  de 0,99 et une erreur quadratique moyenne (RMSE) de  $38,70 \text{ Wh/m}^2$  pour la phase totale.

La deuxième contribution consiste à développer un modèle de bootstrap réseaux de neurones agrégés (BRNA) valable pour la prédiction de l'irradiation solaire horaire sur la ville de Bouzaréah. La base de données comprenant plusieurs Paramètres météorologiques mesurés a été fournie par Centre de Développement des Énergies Renouvelables (CDER) Bouzaréah– Algérie. Les résultats obtenus sont très satisfaisants et très précis avec un coefficient de corrélation de 0.9680, une racine erreur quadratique moyenne (REQM) de  $62,4856 \text{ Wh/m}^2$  et une erreur standard de prédiction (ESP) de 12.79 %. Les résultats obtenus indiquent que le modèle bootstrap réseaux de neurones agrégés (BRNA) est mieux adapté et plus résistant pour prédire l'irradiation solaire globale horaire reçue sur une surface horizontale.

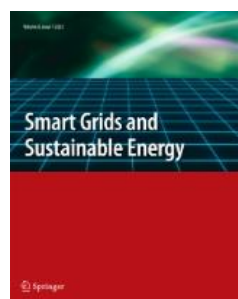
**Mots-clés :** Intelligence artificielle ; Modélisation; Irradiation solaire ; Etude comparative.

## Production Scientifique

### Liste de publications scientifiques



**A. Dahmani**, Y. Ammi, N. Bailek, A. Kuriqi, N. Al-Ansari, S. Hanini, I. Colak, L. Abualigah, S. El-kenawy “Assessing the Efficacy of Improved Learning in Hourly Global Irradiance Prediction” *Computers, Materials & Continua* (2023) vol.77, no.2. <https://doi.org/10.32604/cmc.2023.040625>



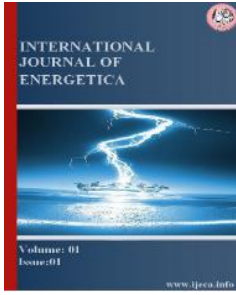
**A. Dahmani**, Y. Ammi, S. Hanini, “A Novel Non-Linear Model Based on Bootstrapped Aggregated Support Vector Machine for the Prediction of Hourly Global Solar Radiation” *Smart Grids and Sustainable Energy* (2024) 9:3. <https://doi.org/10.1007/s40866-023-00179-w>



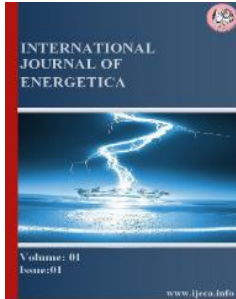
**A. Dahmani**, Y. Ammi, S. Hanini, M. Redha Yaiche, and H. Zentou, “Prediction of Hourly Global Solar Radiation: Comparison of Neural Networks/Bootstrap Aggregating,” *Kem. u Ind. Časopis kemičara i Kem. inženjera Hrvat.*, vol. 72, no. 3–4, pp. 201–213, 2023. <https://doi.org/10.15255/KUI.2022.065>



I. Koudri, D. Kaidameur, **A. Dahmani**, R. Al-Sabur, A. Benyekhlef, A. Sharkawy, “Comparative analysis of fouling resistance prediction in shell and tube heat exchangers using advanced machine learning techniques” *Research on Engineering Structures and Materials.*, pp.18,2023 <https://dx.doi.org/10.17515/resm2023.858en0816>



**A. Dahmani**, Y. Ammi, S. Hanini “Neural network for prediction solar radiation in Relizane region(Algeria) -Analysis study,” Int. J. Energetica., vol.07, no. 2, pp. 08-18, 2022. <https://www.ijeca.info>.



**A. Dahmani**, Y. Ammi, S. Hanini, Z. Driss “Developed nonlinear model based on bootstrap aggregated neural networks for predicting global hourly scale horizontal irradiance” Int. J. Energetica., vol.08, no. 3, pp. 31-37, 2023. <https://www.ijeca.info>.

### Liste de communications

**A. Dahmani**, Y. Ammi, S. Hanini, “Using Artificial Neural Networks to Predict Solar Radiation for Biskra City, Algeria”. Algerian Symposium on Renewable Energy and Materials (ASREM), Médéa (Algeria). 16-17 mars 2020.

**A. Dahmani**, Y. Ammi, S. Hanini, I. Kouidri “Assessment of Neural Networks and Multiple Linear Regressions Models to Predict Solar Irradiance”. National Seminar of Physics, Chemistry and Their Applications (NSPCA’23), Mohamed El Bachir El Ibrahimi University, Bordj-Bou-Arredj, March 6<sup>th</sup>-7<sup>th</sup>, 2023.

**A. Dahmani**, Y. Ammi, S. Hanini, I. Kouidri “Bootstrap Aggregated Neural Networks for Predicting Global Solar Radiation”. Algerian Conference on Mechanics and Solar Energy (ACMSE-2023), Ouargla University, Ouargla, June.7<sup>th</sup>-8<sup>th</sup>, 2023.

**A. Dahmani**, Y. Ammi, S. Hanini, “Deep Learning and Machine Learning Technique for Predicting tilted solar radiation”. 1st National Conference On Science & Technology(1<sup>st</sup> NCST2022), Mascara. 27<sup>th</sup>-28<sup>th</sup> June 2022

**A. Dahmani**, Y. Ammi, S. Hanini, “The Optimized Artificial Neural Network Model With Levenberg-Marquardt Algorithm For Global Solar Radiation Estimation In Ouargla City Of Algerai”. ahi evran international conference on scientific research, Kırşehir Ahi Evran University. Turkey. November 30 - December 1-2, 2021.



**A. Dahmani**, Y. Ammi, S. Hanini, “Prediction Of 5-Min Time-Step Data Of Horizontal Solar Global Irradiation Using Ann (Artificial Neural Network) Model”. 7th International Zeugma Conference On Scientific Researches, Gaziantep University. Turkey. January 21-23, 2022.

**A. Dahmani**, Y. Ammi, S. Hanini, “Supervised Artificial Neural Network-Based Method For Prediction of Solar Radiation Data: Case Study”. ISPEC 8th International Conference On Agriculture, Animal Sciences and Rural Development, Bingol, Turkey. December 24-25, 2021.

**A. Dahmani**, Y. Ammi, S. Hanini, “Artificial intelligence models for predict the hourly global solar radiation of different algorithms training and transfer function in medea of Algeria”. ISPEC 9<sup>th</sup> International Conference On Agriculture, Animal Science And Rural Development, Burdur, Turkey. March 19-20, 2022.

**A. Dahmani**, Y. Ammi, S. Hanini, “A Comparative Assessment Of Multi-Models Based Artificial Intelligence To Predict Global Solar Irradiance In Horizontal Plane”. The Ankara International Congress On Scientific Research, Ankara, Turkey. April 1-3, 2022.

**A. Dahmani**, Y. Ammi, S. Hanini, “An Artificial Intelligence Method For Predicting Solar Radiation Based Neural Networks And Support Vector Machine”. 8<sup>th</sup> International Mardin Artuklu Scientific Researches Conference, Mardin, Turkey. June 04-06, 2022.

## Table des matières

Dédicaces	I
Remerciement	II
Résumés	III
Production Scientifique	VI
Table des matières	IX
Nomenclatures	XII
Liste des tableaux	XV
Liste des figures	XVI

<b>Introduction générale</b>	<b>1</b>
------------------------------	----------

### **Chapitre I: Généralités et notions des bases sur l'irradiation solaire**

I.1. Introduction	4
I.2. Source solaire	4
I.2.1. Structure de soleil	5
I.3. Aspect géométrique - Position du Soleil	6
I.3.1. Mouvement du globe terrestre et la distance terre-soleil	6
I.4. Irradiation solaire	7
I.5. Coordonnées géographiques terrestres	9
I.6. Systèmes de coordonnées, équatoriales et horizontales	9
I.7. Impact de l'atmosphère sur l'irradiation solaire	12
I.7.1. Atmosphère	12
I.7.2. Composition de l'atmosphère terrestre	12
I.7.3. Masse atmosphérique	13
I.7.4. Indice de clarté	14
I.8. Mesure d'irradiation solaire	14
I.8.1. Radiomètres solaires	14
I.9. Mesures Satellites météorologiques	16
I.10. Rayonnement solaire dans le monde	17
I.10.1 Potentiel de l'énergie solaire en Algérie	19
I.10.2. Programme de développement des énergies renouvelables en Algérie	21
I.11. Conclusion	22

### **Chapitre II: Synthèse bibliographique sur la modélisation de l'irradiation solaire par les modèles d'intelligence artificielle**

II.1. Introduction	24
II.2. Production scientifique mondiale sur la prédiction de l'irradiation solaire	24
II.3. Intelligence artificielle	25
II.3.1. Réseaux de neurones artificiels (RNA)	26

II.3.1.1. Neurone biologique	26
II.3.1.2. Neurone artificiel (formel)	26
II.3.1.3. Fonctions d'activation	28
II.3.1.4. Architecture d'un réseau de neurones	28
II.3.1.5. Perceptron multicouches	29
II.3.1.6. Apprentissage	30
II.3.1.7. Sur-apprentissage et sous-apprentissage	31
II.3.1.8. Machine D'apprentissage Extrême (ELM)	31
II.3.2. Machines à vecteurs support (MVS)	33
II.3.3. Forêt aléatoire (FA)	35
II.3.4. Gradient Boosting Extrême (XGBoost)	36
II.4. Algorithmes d'analyse de séries temporelles	37
II.4.1. Autorégressif non linéaire avec entrées exogènes (ARNX)	37
II.4.2. Autorégressif intégré moyenne mobile (ARIMA)	38
II.5. Revue des travaux sur la modélisation de l'irradiation solaire	41
II.6. Performance et évaluation des prédicteurs	49
II.7. Conclusion	50

### **Chapitre III: Évaluation de l'efficacité de l'apprentissage amélioré dans la prédiction de l'irradiation horizontale globale**

III.1. Introduction	51
III.2. Région étudiée et collecte de données	52
III.3. Evaluation et développement du modèle	53
III.4. Analyse de sensibilité par la méthode de "poids"	55
III.5. Diagnostic des points aberrants et domaine d'applicabilité du modèle neuronal optimisé	57
III.6. Résultats et discussion	58
III.6.1. Expressions mathématique des modèles RNA-PMC	63
III.6.2. Analyse de sensibilité	68
III.6.3. Diagnostic des valeurs aberrantes et domaine applicabilité des modèles optimaux	70
III.7. Conclusion	72

### **Chapitre IV: Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux / agrégation bootstrap**

IV.1. Introduction	74
IV.2. Collecte des données météorologiques, prétraitement et analyse.	75
IV.3. Procédure de modélisation	75
IV.3.1. Modélisation avec des réseaux neuronaux (RNU et RNI)	75
IV.3.2. Modélisation avec des bootstrap réseaux de neurones agrégés (BRNA)	77
IV.4. Résultats et discussion	78
IV.4.1. Effet de division de la base des données	78
IV.4.2. Modèles de performance	80
IV.4.3. Comparaison entre RNI, RNU et BRNA	84
IV.4.4. Analyse de sensibilité	92

IV.4.5. Diagnostic des points aberrants et domaine d'applicabilité du modèle neuronal optimisé	94
IV.4.6 Comparaison avec d'autres modèles	96
IV.5. Conclusion	97
<b>Conclusion générale</b>	99
<b>Références bibliographiques</b>	102

## Nomenclatures

Symbole	Signification	Unité
<b>Abréviations</b>		
IA	Intelligence Artificielle	--
MVS	Machines à Vecteur de Support	--
RNA	Réseaux de Neurones Artificiels	--
PMC	Perceptron à Multicouches	--
RBF	Fonction de Base Radiale	--
BRNA	Bootstrap Réseaux de Neurones Agrégés	--
RNU	Réseaux Neurones Unique	--
RNI	Réseaux Neurones Individuels	--
AG	Algorithme Génétique	--
HS	Systèmes Hybrides	--
ELM	Extreme Learning Machine	--
MVSR	Machines à Vecteurs Support pour la Régression	--
FA	Forêt Aléatoire	--
XGBoost	Extreme Gradient Boosting	--
ARNX	Autorégressif non Linéaire avec Entrées Exogènes	--
ARIMA	Autorégressif Intégré Moyenne Mobile	--
BVAR	Bayesian Vecteur Auto Régression	--
RNFF	Réseau neuronal Feed Forward	--
ANFIS	Systèmes d'Inférence Neuro-Flous Adaptatifs	--
GRNN	Réseau Neuronal de Régression Généralisée	--
RNB	Réseau Neuronal Bayésien	--
NASA	National Aéronautique and Space Administration	--
CDER	Centre de Développement des Énergies Renouvelables	--
CIB	Critère d'Information bayésien	--
CIA	Critère d'Information d'Akaike	--
BFGS	Broyden–Fletcher–Goldfarb–Shanno	--

WMO	Organisation Météorologique Mondiale	--
ECM	Error Correction Models	--
TSV	Temps Solaire Vrai	min
REQM	Racine Erreur Quadratique Moyenne	W/m <sup>2</sup>
ESP	Erreur Standard de Prédiction	%
EAM	Erreur Absolue Moyenne	W/m <sup>2</sup>
EBM	Erreur de Biais Moyenne	%
EMAP	Erreur Moyenne Absolue en Pourcentage	%
EnR	Énergies Renouvelables	--
ASA	Agence Spatiale Allemande	--
CSP	Énergie Solaire Concentrée	--
PV	Photovoltaïque	--
ISG	Irradiation Solaire Globale	W/m <sup>2</sup>
UV	Ultraviolet	nm
IR	Infrarouge	µm
VMD	Vitesse Moyenne de vent	m/s
DV	Direction de Vent	°
MA	Masse Atmosphérique	m

### Lettres latines

H <sub>F</sub>	Irradiation Diffuse	W/m <sup>2</sup>
H <sub>D</sub>	Irradiation Solaire Direct	W/m <sup>2</sup>
H	Irradiation Solaire Globale	W/m <sup>2</sup>
h	Hauteur du soleil	m
K <sub>T</sub>	Indice de clarté	--
f	Fonction	--
w	Poids	--
<i>b<sub>j</sub></i>	Biais	--
Tansig	Fonction D'activation Tangente Hyperbolique	--
Logsig	Sigmoïde Logarithmique	--
Purelin	Fonction d'Activation Linéaire	--
Sin	Fonction d'Activation Sinus	--
Exponentiel	Fonction d'Activation Exponentielle	--
R	Coefficient de corrélation	%

$R^2$	Coefficient de détermination	%
IHG	Irradiation Horizontale Globale	W/m <sup>2</sup>
BD	Base de données	--
IR	Importance relative	%
H*	valeur critique de levier	--

### Indices

i	Couche d'entrée	--
h	Couche cachée	--
o	Couche de sortie	--
min	Minimal	--
max	Maximal	--
DS	Déviatiion Standard	--

### Lettres grecques

$\theta_z$	Angle zénithal.	°
$\theta_i$	Angle d'incidence	°
$\theta$	Latitude	°
$\varphi$	Longitude	°
$\delta$	déclinaison	°
$\psi$	L'azimut solaire	°
$\omega$	Angle horaire	°

## Liste des tableaux

### Chapitre I : Généralités et notions des bases sur l'irradiation solaire

<b>Tableau I.1.</b> Distribution de l'énergie solaire en Algérie.	21
<b>Tableau I.2.</b> Consistance du programme de développement des EnR.	22

### Chapitre II : Synthèse bibliographique sur la modélisation de l'irradiation solaire par les modèles d'intelligence artificielle

<b>Tableau II.1.</b> Fonctions de transfert usuelles.	28
<b>Tableau II.2.</b> Exemples significatifs de l'application des différents types publiés à la prédiction de l'irradiation solaire.	48

### Chapitre III : Évaluation de l'efficacité de l'apprentissage amélioré dans la prédiction de l'irradiation horizontale globale

<b>Tableau III.1.</b> Région géographique et période couverte par les deux stations.	52
<b>Tableau III.2.</b> Analyse statistique des variables d'entrée et de sortie pour la prédiction de l'irradiation horizontale globale à deux stations.	53
<b>Tableau III.3.</b> Corrélations climat-sortie pour la prédiction de l'irradiation horizontale globale à deux stations.	53
<b>Tableau III.4.</b> Évaluation statistique des erreurs du modèle (RNA-PMC) sur la prédiction de l'irradiation horizontale globale horaire.	59
<b>Tableau III.5.</b> Structure du réseau de neurones optimisé.	63
<b>Tableau III.6.</b> Évaluation des performances des modèles MVS avec différents Fonction noyaux.	65
<b>Tableau III.7.</b> Évaluation les erreurs des modèles élaborés de prédiction de l'irradiation horizontal horaire pour les deux stations.	68
<b>Tableau III. 8.</b> Poids et biais du RNA-PMC optimal.	70

### Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux / agrégation bootstrap

<b>Tableau VI.1.</b> Analyse statistique des données d'entrée et de sortie.	76
<b>Tableau IV.2.</b> Influence de la division de la base de données pour RNU.	80
<b>Tableau IV.3.</b> Structures et performances des modèles RNI.	81
<b>Tableau IV.4.</b> Performances des modèles RNI, RNU et BNRA en termes de R, EAM, EMP, REQM et ESP.	85
<b>Tableau IV. 5.</b> Poids et biais de réseau neurone unique.	93
<b>Tableau IV.6.</b> Comparaison des résultats de notre modèle avec différents modèles similaires.	97



# Liste des figures

## Chapitre I : Généralités et notions des bases sur l'irradiation solaire

<b>Figure I.1.</b> Structure du soleil.	5
<b>Figure I.2.</b> Mouvement de révolution de la Terre autour du Soleil.	6
<b>Figure I.3.</b> Variation de l'excentricité et de la distance terre-soleil.	7
<b>Figure I.4.</b> Distribution spectrale du rayonnement solaire.	8
<b>Figure I.5.</b> Coordonnées terrestres d'un lieu donné.	9
<b>Figure I.6.</b> Variation saisonnière de la déclinaison solaire.	11
<b>Figure I.7.</b> Système de coordonnées solaires dans le plan horizontal.	12
<b>Figure I.8.</b> a) Schéma d'un pyréliomètre. b) Photo de Hukseflux DR01.	14
<b>Figure I.9.</b> a) Schéma d'un pyranomètre. b) Pyranomètre type EKO MS-410.	15
<b>Figure I.10.</b> a) Schéma d'héliographe Campbell-Stokes. b) Photo typique d'héliographe Campbell-Stokes).	16
<b>Figure I.11.</b> Capteur Albédomètre.	16
<b>Figure I.12.</b> Intensité de l'irradiation solaire mondiale.	18
<b>Figure I.13.</b> Durée d'insolation en Algérie.	20
<b>Figure I.14.</b> Présentation des ratios du potentiel solaire.	20

## Chapitre II : Synthèse bibliographique sur la modélisation de l'irradiation solaire par les modèles d'intelligence artificielle

<b>Figure II.1.</b> Nombre de publications inhérents à la prédiction d'irradiation solaire publiés durant les années avant 1994 à 2023.	25
<b>Figure II.2.</b> Représentation d'un neurone biologique.	26
<b>Figure II.3.</b> Représentation mathématique d'un neurone artificiel.	27
<b>Figure II.4.</b> Structure d'un réseau de neurones non bouclé.	30
<b>Figure II.5.</b> Compromis biais-variance d'un modèle optimisée.	31
<b>Figure II.6.</b> Structure d'un modèle ELM	32
<b>Figure II.7.</b> Exemple de projection	33
<b>Figure II.8.</b> Différents architecture du réseau NARX	38

## Chapitre III : Évaluation de l'efficacité de l'apprentissage amélioré dans la prédiction de l'irradiation horizontale globale

<b>Figure III.1.</b> Diagramme de la méthodologie pour calculer les deux modèles (RNA-PMC) et (MVS).	54
<b>Figure III.2.</b> Diagramme alternatif de la méthode de poids.	56
<b>Figure III.3.</b> Effet de la division de la base de données en corrélation des coefficients (R) pour la phase de test.	59

<b>Figure III.4.</b> Modèle les valeurs expérimentales vs. Les valeurs calculées pour les phases de données : (a) phase d'apprentissage, (b) phase de validation, (c) phase de test, (d) totale.	60
<b>Figure III.5.</b> Architecture du réseau neuronal.	63
<b>Figure III.6.</b> Représentation schématique du RNA-PMC amélioré.	64
<b>Figure III.7.</b> Modèle MVS-RBF les valeurs expérimentales vs. Les valeurs calculées pour les phases de données : (a) phase d'apprentissage, (b) phase de test, (c) totale.	66
<b>Figure III.8.</b> Présentation de L'importance relative (RI) des paramètres d'entrées sur la prédiction de l'irradiation horizontale globale.	69
<b>Figure III.9.</b> Diagrammes de Williams pour le modèle MVS-RBF optimal de prédiction de l'irradiation horizontale globale horaire.	71
<b>Figure III.10.</b> Diagrammes de Williams pour le modèle RNA-PMC optimal de prédiction de l'irradiation horizontale globale horaire.	72

#### **Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux / agrégation bootstrap**

<b>Figure IV.1.</b> Localisation de la station météorologique CDER à Bouzaréah.	75
<b>Figure IV.2.</b> Diagramme de flux pour le développement de réseaux neuronaux (RNU RNI et BRNA <sub>(stacking de 30 réseaux)</sub> ).	76
<b>Figure IV.3.</b> Réseaux de neurones agrégés bootstrap.	78
<b>Figure IV.4.</b> Irradiation solaire global Vs température.	79
<b>Figure IV.5.</b> Comparaison des valeurs expérimentales Vs. valeurs calculées : (a) la phase de validation(RNU), (b) la phase de test (RNU), (c) la phase de validation (RNAB), (d) la phase de test (RNAB).	82
<b>Figure IV.6.</b> EPM, REQM, et ESP pour BRNA et RNU : (a) phase de validation, (b) phase de test.	92
<b>Figure IV.7.</b> Illustration d'importance relative (IR) de chaque variable d'entrée par rapport à la sortie.	93
<b>Figure IV.8.</b> Diagnostic des données potentiellement suspectes et domaine d'applicabilité du modèle neuronal optimal "RNA-PMC" pour : (a) phase de validation, (b) phase de test.	94
<b>Figure IV.9.</b> Diagnostic de données potentiellement suspectes et délimitation du domaine d'applicabilité du modèle neuronal "BRNA" pour : (a) phase de validation, (b) phase de test.	96

# Introduction générale

De nos jours, le monde connaît une croissance démographique rapide qui s'accompagne d'une amélioration de la qualité de vie, laquelle repose en grande partie sur l'accès à l'électricité. Cette croissance présente en effet de nombreux défis pour l'humanité. De nombreux habitants des pays en développement aspirent à un mode de vie similaire à celui des pays riches, ce qui a entraîné une forte augmentation de la demande mondiale en énergie, alors même que nos réserves de combustibles fossiles diminuent. Selon les données du ministère de l'énergie algérien, la consommation de gaz naturel devrait passer de 42 milliards de mètres cubes à 75 milliards de mètres cubes d'ici 2030 [1]. Dans ce contexte, il est devenu essentiel d'accélérer la transition vers une nouvelle politique énergétique afin de ralentir l'épuisement de nos réserves en hydrocarbures, tout en continuant à répondre aux besoins de confort et de consommation des populations.

Dans ce contexte, de nombreux pays à travers le monde envisagent de diversifier leurs sources d'approvisionnement en énergie en intégrant de nouvelles ressources dans leurs réseaux électriques, notamment les énergies renouvelables. L'énergie solaire se positionne comme la ressource la plus abondante et durable sur terre. L'Algérie se distingue par l'un des gisements solaires les plus vastes au monde, bénéficiant d'une durée d'insolation dépassant les 2000 heures par an sur son territoire. Cela confère à la mise en place de projets de production d'énergie solaire un potentiel de rendement exceptionnel. Parmi les principales applications énergétiques utilisant l'éclairage solaire, on compte le photovoltaïque, le chauffage et le refroidissement solaire, ainsi que l'énergie solaire à concentration.

Un historique de données s'avère essentiel pour plusieurs étapes cruciales telles que la sélection du site, l'étude de faisabilité, le dimensionnement optimisé et les analyses technico-économiques d'un projet. Étant donné que les conditions météorologiques peuvent varier d'une année à l'autre, il est impératif que le promoteur de projet ait accès à des données historiques couvrant différentes saisons, jours, heures voire de préférence des sous-périodes, permettant ainsi de prédire les performances annuelles et quotidiennes d'une centrale électrique envisagée.

Lors de temps de mise en service de l'unité de production d'énergie, des mesures précises en temps réel se révèlent extrêmement bénéfiques. Elles servent à effectuer des tests de performance approfondis et à surveiller en continu le rendement de l'installation. Cette surveillance constante garantit l'efficacité de l'installation tout au long de sa durée de vie.

L'irradiation solaire est la principale source d'énergie disponible à la surface de la terre. Il est utile pour de nombreux processus physiques et chimiques. De nombreux processus biologiques ne peuvent

avoir lieu sans l'irradiation solaire. En plus de ces processus, il est également utilisé directement et indirectement pour la production d'électricité. En raison de l'augmentation rapide de la demande en énergie à l'échelle mondiale, il existe une recherche mondiale d'une source alternative de production d'énergie. Cela est dû aux effets destructeurs des sources d'énergie non renouvelables sur l'environnement. Par conséquent, il y a une demande croissante de données sur l'irradiation solaire pour le déploiement rapide de technologies d'énergie propre et verte tels que les systèmes photovoltaïques dans le monde entier [2]. Les données de l'irradiation solaire jouent un rôle significatif dans la conception des systèmes d'énergie solaire. Cependant, Elles ne sont pas largement accessibles en raison du nombre restreint de stations météorologiques. La meilleure méthode pour obtenir l'irradiation solaire globale (ISG) en tout lieu consiste à installer des capteurs (pyranomètres, pyréliomètres, etc.) à chaque emplacement spécifique. Cependant, en raison des difficultés d'enregistrement, de calibrage et d'entretien, cela représente une tâche laborieuse et coûteuse [3], [4]. Dans les pays en développement, la mesure fiable des données de ISG est relativement rare en raison des raisons évoquées ci-dessus [3].

Cette situation a encouragé l'amélioration des méthodes de calcul comme la modélisation. La modélisation de l'irradiation solaire joue un rôle crucial dans le domaine de l'énergie solaire. La prédiction de l'irradiation solaire se révèle particulièrement complexe en raison de sa dépendance à l'égard de multiples paramètres géographiques et astronomiques, tout en étant influencée par les variations des conditions astronomiques et météorologiques. Par conséquent, il devient impérieux de concevoir des modèles qui reposent sur les données recueillies par les stations disponibles, englobant des éléments comme l'irradiation extraterrestre, durée d'ensoleillement, température, précipitations l'humidité, pression atmosphérique, la vitesse et direction du vent, latitude, l'altitude, longitude et le taux de couverture nuageuse.

Au fil du temps, plusieurs méthodes ont été développées pour prédire l'irradiation solaire à la surface de la terre. Parmi celles-ci, on peut citer les modèles empiriques [5] ainsi qu'un modèle hybride Évolutif [6]. Cependant, ces approches ont démontré des performances limitées lorsqu'il s'agit de prévoir les données de l'irradiation solaire à long terme. Dans le but de surmonter ces défis, les recherches récentes se sont tournées vers l'utilisation de techniques d'intelligence artificielle basées sur les réseaux de neurones artificiels et machine à vecteur support (MVS) [7].

Dans ce contexte que s'inscrit l'objectif de notre thèse. Notre travail se concentre sur l'exploration de ces nouvelles approches afin d'améliorer la prédiction de l'irradiation solaire horaire. En utilisant des réseaux de neurones uniques, bootstrap réseaux neurones agrégés et machine à vecteur support (MVS),

nous visons à fournir des prédictions plus précises et fiables de l'irradiation solaire horaire, ouvrant ainsi la voie à des progrès significatifs dans ce domaine.

Cette thèse s'organise en 5 chapitres:

- Le premier chapitre est consacré à la description générale des notions principales sur le soleil et l'irradiation solaire, On étudie les divers paramètres astronomiques et géographiques susceptibles d'exercer une influence sur l'intensité de l'irradiation solaire reçue à la surface terrestre, ainsi que le potentiel solaire à l'échelle mondiale, particulièrement en Algérie.
- Dans le deuxième chapitre, on décrit les techniques d'intelligence artificielle et détaille synthèse bibliographique sur la prédiction de l'irradiation, et les objectifs poursuivis par cette étude.
- Le troisième chapitre exposera première contribution pour précision de la prédiction de l'irradiation horizontale globale horaire sur les deux sites Tamanrasset et Bouzaréah basée sur les paramètres météorologiques, en utilisant des modèles conventionnels d'apprentissage automatique, en particulier les réseaux neurone artificiel type de perceptron multicouches (RNA-PMC) et machine à vecteur support (MVS).
- Ensuite, dans le quatrième chapitre, nous présentons la deuxième contribution était de développer un modèle bootstrap réseaux neurones agrégés (BRNA) qui utilisé pour prédire l'irradiation solaire horaire sur le site de Bouzaréah. Il confirme qu'un tel modèle de bootstrap réseaux neuronaux agrégés (BRNA) peut donner une prédiction plus précise et plus robuste qu'un modèle de réseaux neuronaux unique (RNU).

Enfin, La thèse se clôturera par une conclusion générale qui synthétisera les résultats obtenus au cours de cette recherche ainsi que les perspectives envisageables pour l'avenir.

# **Chapitre I**

Généralités et notions des bases sur l'irradiation solaire

### I.1. Introduction

Pour une utilisation optimale de l'énergie solaire et pour maximiser sa capture à travers divers systèmes énergétiques, il est essentiel de comprendre en profondeur la répartition de l'irradiation solaire. L'irradiation solaire émise par le soleil dans tout l'espace se présente sous la forme d'ondes électromagnétiques qui transportent de l'énergie à la vitesse de la lumière. Le rayonnement solaire est absorbé, réfléchi ou diffusé par les particules solides en tout point de l'espace, en particulier par la Terre. Ce processus dépend de nombreuses conditions environnementales telles que la météo, le climat, la pollution, etc. Le rayonnement entrant en un point donné prend différentes formes en fonction de son emplacement géographique, de ses coordonnées astronomiques, de sa distance par rapport au soleil, de la composition de l'atmosphère locale et de la topographie locale. Ce chapitre couvre quelques concepts de base, définitions et équations astronomiques, ainsi qu'une description des capteurs radiométriques et de leurs principes de fonctionnement. Enfin, nous consacrons une brève discussion au paysage mondial de l'énergie solaire. Cette section fournit des informations utilisées dans notre thèse.

### I.2. Source solaire

Le soleil est la principale source d'énergie sur Terre, fournissant la puissance nécessaire à la photosynthèse et provoquant des mouvements d'air et d'eau [8]. La constante solaire représente la quantité d'énergie qui atteint une surface faisant face au soleil ; elle équivaut à 1380 J/s et sa variation moyenne est de 0,2 % [9]. L'irradiation solaire est naturellement exprimée sous forme de lumière et de chaleur, et peut être transformée en électricité grâce à des panneaux solaires. L'énergie solaire peut être capturée grâce à une conception solaire passive (comme de grandes fenêtres), à des systèmes thermiques (pour chauffer l'eau) ou être récoltée par des panneaux solaires (pour générer de l'électricité). L'électricité peut être stockée ou utilisée pour chauffer et éclairer les bâtiments [10].

Suri et Hofierka [11] décrivent trois catégories de facteurs affectant l'énergie solaire : les caractéristiques de la Terre, les caractéristiques du terrain et l'extinction atmosphérique.

Les caractéristiques de la Terre déterminent l'irradiation solaire extraterrestre, c'est-à-dire l'énergie disponible en haut de l'atmosphère. Cette énergie varie en fonction de la position relative du Soleil par rapport à l'horizon et est influencée par la géométrie de la Terre et ses mouvements (rotation et translation). La position exacte du Soleil par rapport à un point sur Terre peut être calculée avec précision en utilisant des formules astronomiques. Elle dépend de la déclinaison de la Terre (saison), de la latitude et de l'angle horaire solaire (heure de la journée) [9].



La topographie influence la quantité de l'irradiation solaire disponible dans un endroit spécifique. Elle dépend des caractéristiques du paysage, notamment l'altitude, la pente, l'orientation et l'ombre. L'atténuation de l'énergie due à ces facteurs peut être précisément modélisée à partir de données topographiques.

L'atténuation atmosphérique est causée par la diffusion et l'absorption, provoquées à la fois par les gaz et les particules solides et liquides en suspension (y compris les nuages) dans l'atmosphère. L'atténuation atmosphérique implique des interactions complexes et peut être modélisée avec un niveau de précision modéré.

L'extinction atmosphérique réduit l'énergie solaire disponible en un endroit spécifique en raison de l'absorption et de la diffusion provoquées par les composants atmosphériques, qu'il s'agisse de solides, de liquides ou de gaz (y compris les nuages). Cette catégorie dépend des composants atmosphériques et de l'altitude, qui déterminent l'épaisseur de l'atmosphère [9].

### I.2.1. Structure de soleil

La structure du soleil est illustrée dans la figure I.1 qui se divise en quatre régions principales : le noyau, la photosphère, la chromosphère et la couronne. En conséquence, le rayonnement solaire émis dans l'espace est le produit d'une interaction complexe entre ces différentes couches, chacune émettant et absorbant des radiations à diverses longueurs d'onde [12].

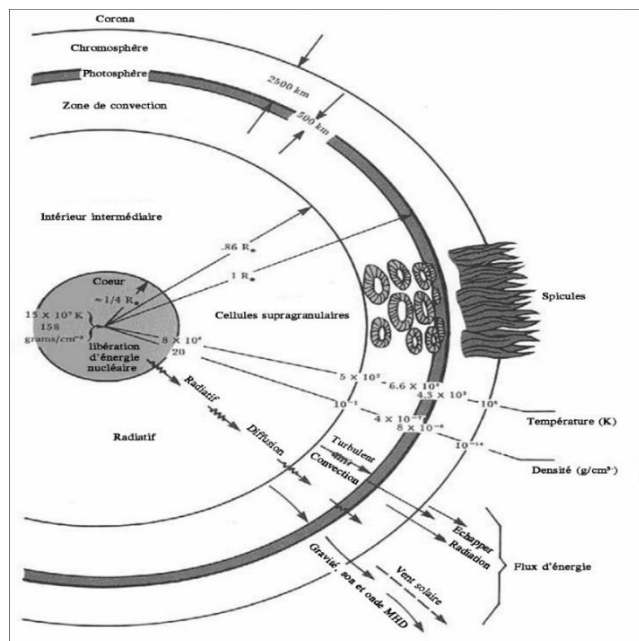


Figure I.1. Structure du soleil [13].

### I.3. Aspect géométrique - Position du Soleil

Dans ce qui suit, nous envisagerons les paramètres astronomiques de base et leur la définition du calcul du rayonnement solaire. Ces concepts de base, définitions et équations astronomiques dérivées fournissent la base pour évaluer l'énergie solaire à un endroit donné.

#### I.3.1. Mouvement du globe terrestre et la distance terre-soleil

La terre parcourt une orbite elliptique autour du soleil, avec le soleil situé en l'un de ses points focaux, sur une période de 365,25 jours. Le plan dans lequel la terre effectue cette orbite est appelé plan de l'écliptique. Le mouvement de la terre autour du soleil se fait à une vitesse moyenne de km/s (atteignant un maximum en hiver à 30,27 km/s et un minimum en été à 29,77 km/s). Ce mouvement s'effectue dans le sens trigonométrique direct et engendre le cycle des saisons (Figure I.2). La terre est au point le plus proche du soleil vers le 3 janvier et au plus éloignée vers le 4 juillet. Un jour solaire est défini comme la durée moyenne entre deux passages successifs du soleil au méridien d'un endroit donné, divisée en 24 segments égaux appelés heures. [14], La quantité de rayonnement solaire atteignant la terre est inversement proportionnelle au carré de la distance séparant la terre du soleil.

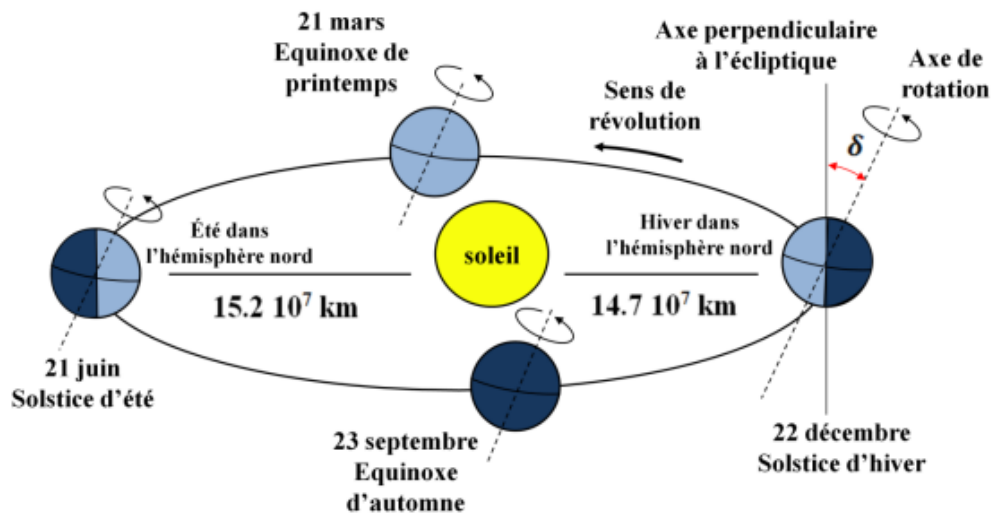


Figure I.2. Mouvement de révolution de la Terre autour du Soleil [1].

La figure I.3 illustre la variation de l'excentricité de l'orbite terrestre ainsi que de la distance entre la Terre et le Soleil au cours de l'année.

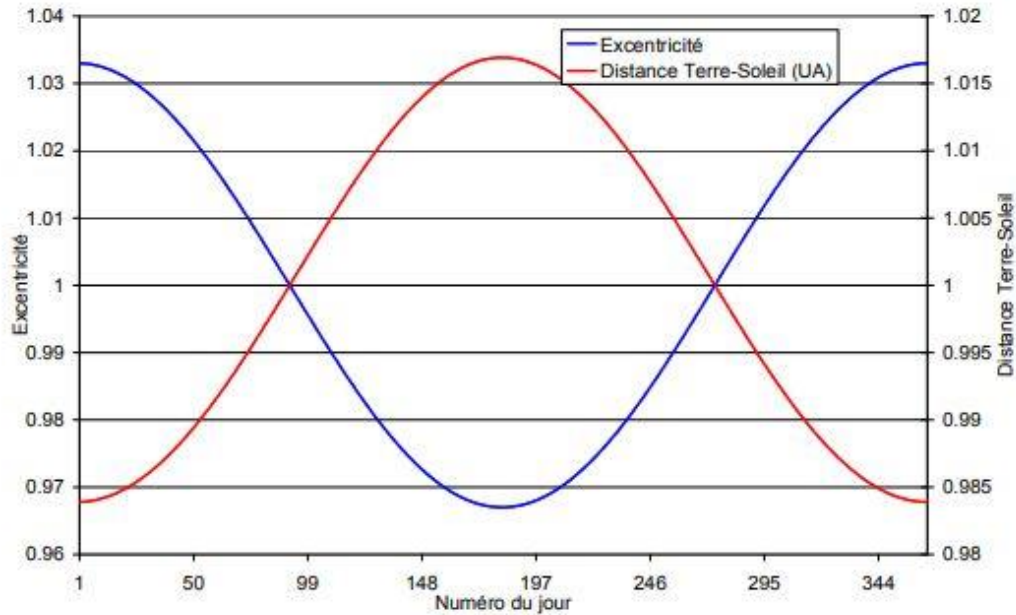


Figure I.3. Variation de l'excentricité et de la distance terre-soleil [15].

### I.4. Irradiation solaire

L'irradiation solaire est l'irradiation émise par le soleil par des réactions thermonucléaires au sein de son noyau. Le rayonnement solaire voyage dans l'espace sous forme d'ondes électromagnétiques avec une large gamme de longueur d'onde, des ondes radio aux rayons gamma. Cependant, 99% de ce rayonnement contient des longueurs d'onde comprises entre 150 nm et 4000 nm. La répartition spectrale de l'énergie émise par le soleil est la suivante : 40 % en lumière visible, 8 % en ultraviolet (UV) et 51 % en infrarouge (IR). La terre reçoit une partie de l'irradiation [16]. La figure I.4 montre la distribution spectrale du rayonnement solaire au sommet de l'atmosphère, comparée à un corps noir à 5900K et la distribution spectrale reçue sur Terre. L'irradiation reçu en tout point est à la fois direct et diffus.

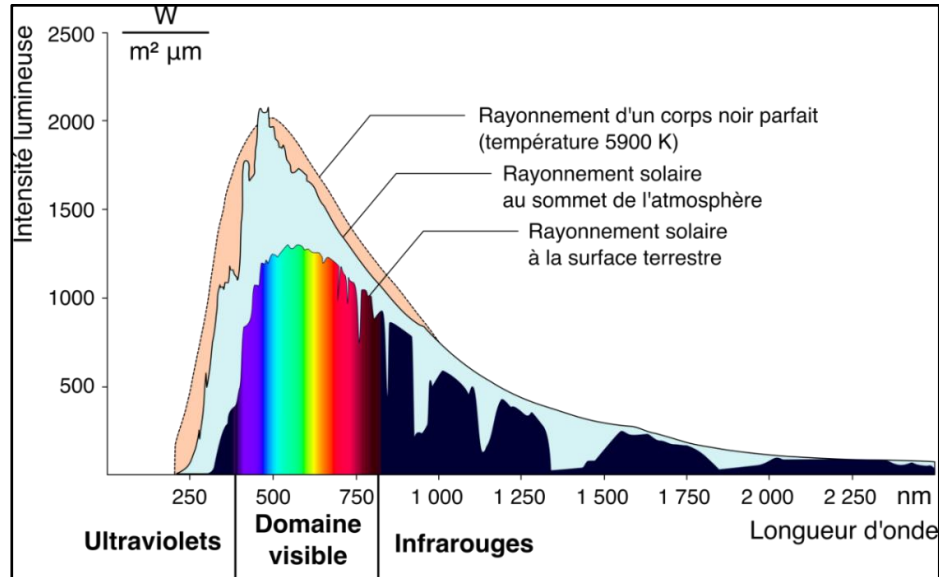


Figure I.4. Distribution spectrale du rayonnement solaire [17].

### Irradiation solaire direct ( $H_D$ )

L'irradiation solaire directe ( $H_D$ ) est définie comme l'irradiation qui se déplace en ligne droite du soleil à la surface de la terre. C'est l'irradiation solaire reçu du soleil sans dispersion par l'atmosphère et sans aucune perturbation. La quantité de l'irradiation solaire direct atteignant n'importe quelle partie particulière de la surface terrestre est déterminée par la position du point, la période de l'année, la forme de la surface [18].

### Irradiation solaire diffus ( $H_F$ )

Une fois que l'irradiation solaire pénètre dans l'atmosphère terrestre, il est partiellement dispersé et partiellement absorbé. L'irradiation diffusé est appelé irradiation diffus ( $H_F$ ). Encore une fois, une partie de cette irradiation diffuse retourne dans l'espace et une partie atteint le sol. L'irradiation solaire diffus est d'abord intercepté par les constituants de l'air tels que la vapeur d'eau, le  $CO_2$ , la poussière, les aérosols, les nuages, etc., puis il est libéré sous forme d'irradiation dispersé dans de nombreuses directions [19].

### Irradiation solaire globale ( $H$ )

La puissance totale du rayonnement solaire reçue à la surface terrestre est calculée en additionnant la puissance de l'irradiation diffus avec la puissance l'irradiation direct, le tout pondéré par le cosinus de l'angle zénithal du soleil [20]. On a alors:

$$H = H_D + H_F \cos(\theta_z) \quad (I.1)$$

$\theta_z$  : l'angle zénithal.

## I.5. Coordonnées géographiques terrestres

Afin de déterminer la position géographique d'un site sur la surface terrestre, plusieurs paramètres sont définis (figure I.5) [21]:

- **Latitude ( $\theta$ )** : Il s'agit de l'angle  $\theta$  formé entre la direction verticale du lieu et le plan équatorial terrestre. Cette mesure varie de 0 à 90 degrés, avec des valeurs positives en se dirigeant vers le pôle nord et des valeurs négatives en se dirigeant vers le pôle sud.
- **Longitude ( $\varphi$ )** : La longitude est également exprimée sous forme d'angle  $\varphi$ , mesuré par rapport de Greenwich et au méridien du lieu spécifique. Cette mesure s'étend de 0 à 180 degrés, avec des valeurs positives en se déplaçant vers l'est et des valeurs négatives en se déplaçant vers l'ouest.
- **Altitude** : L'altitude correspond à la hauteur d'un lieu par rapport au niveau de la mer, et elle est généralement mesurée en mètres (m).

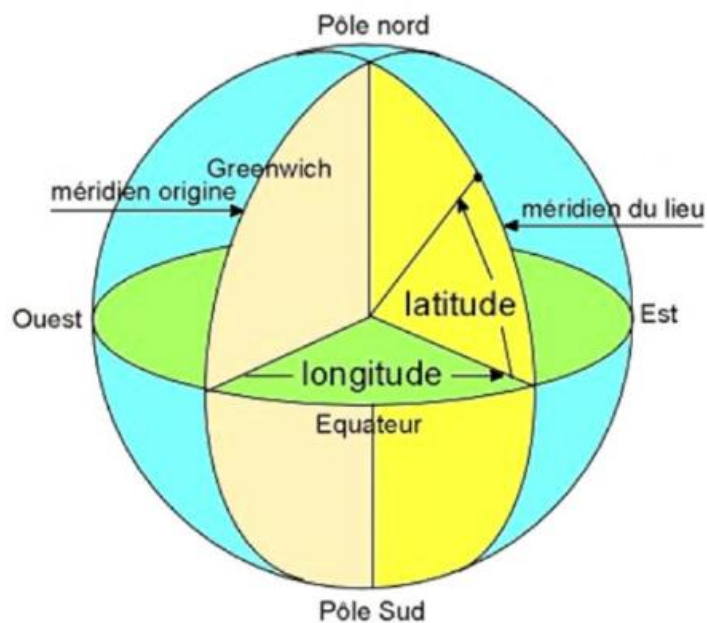


Figure I.5. Les coordonnées terrestres d'un lieu donné [21].

## I.6. Systèmes de coordonnées, équatoriales et horizontales

Relation géométrique entre les plans dans n'importe quelle direction (que ce plan soit immobile ou en mouvement) et le rayonnement solaire incident, c'est-à-dire. La position du soleil par rapport à ce plan peut être décrite par plusieurs angles :

### Angle horaire ( $\omega$ )

L'angle horaire solaire représente l'angle à l'équateur entre le méridien contenant le zénith de l'observateur et le méridien passant par la direction du soleil. Sa valeur est de zéro degrés au moment du midi solaire, ce qui traduit la rotation quotidienne de la Terre autour de son axe polaire. Par conséquent, chaque heure de temps correspond à un angle horaire de 15 degrés, ce qui est donné par la relation suivante : [22]:

$$\omega = (\text{TSV}-12) \times 15^\circ \quad (\text{I.2})$$

$\omega > 0$  : le matin

$\omega < 0$  : l'après-midi

TSV est le temps solaire vrai donné en heures.

Le temps solaire vrai peut être déduit à partir de la relation :

$$\text{TSV} = \text{TU} + \text{ET} + 4\varepsilon |\varphi_r - \varphi| \quad (\text{I.3})$$

$\varphi_r$  est la longitude de référence et  $\varphi$  la Longitude locale

$\varepsilon = +1$  ; pour  $\varphi$  à l'Est de  $\varphi_r$   $\varepsilon = -1$  ; pour  $\varphi$  à l'Ouest de  $\varphi_r$

ET est l'équation du temps (minute) donnée par la relation (I.5) (due à la variation de la vitesse de la Terre autour du Soleil) :

$$\text{ET} = 9.87 \sin(2\gamma) - 7.53 \cos(\gamma) - 1.5 \sin(\gamma) \quad (\text{I.4})$$

Où  $\gamma$  en radian.

TU est le temps universel qui est le temps solaire moyen mesuré à partir de Greenwich (GMT) avec minuit comme origine.

### Déclinaison ( $\delta$ )

La déclinaison représente l'angle entre la direction Terre-Soleil et le plan équatorial de la Terre. L'angle varie entre  $-23,45^\circ$  et  $+23,45^\circ$ . Ces valeurs correspondent respectivement à l'hiver et à l'été (21 décembre et 21 juin), le 21 mars étant annulé (figure I.6).

Plusieurs formule ont été développées pour calculer la déclinaison, la plus simple et précise est donnée par equation suivant [23]:

$$\delta = 23.45 \sin\left(360 \frac{284+n}{365}\right) \quad (\text{I.5})$$

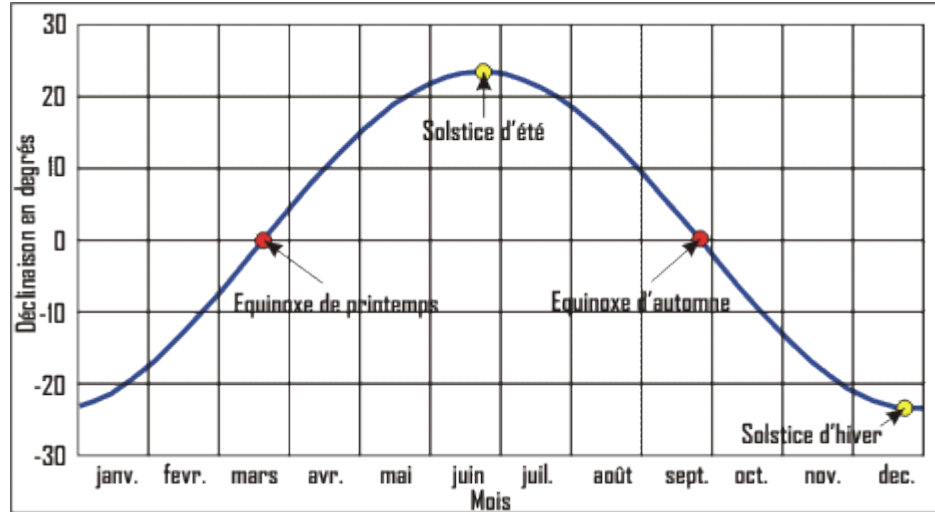


Figure I.6. Variation saisonnière de la déclinaison solaire [12].

### Angle d'incidence ( $\theta_i$ )

Il existe un ensemble précieux de relations entre ces angles. L'équation qui relie l'angle d'incidence sur une surface à d'autres angles est la suivante : [12]:

$$\cos \theta_i = \sin \delta \sin \Phi \cos \beta - \sin \delta \cos \Phi \sin \beta \cos \gamma + \cos \delta \cos \Phi \cos \beta \cos \omega + \cos \delta \sin \Phi \sin \beta \cos \gamma \cos \omega + \cos \delta \sin \beta \sin \gamma \sin \omega \quad (I.6)$$

### Hauteur du soleil (h)

La hauteur du soleil h est l'angle entre les rayons du soleil et le plan horizontal. Elle se compte de 0° à 90° à partir de l'horizon vers la voûte céleste, et varie à chaque instant de la journée et de l'année selon la corrélation suivante [12]:

$$\sin(h) = \cos(\delta)\cos(\Phi)\cos(\omega) + \sin(\Phi)\sin(\delta) \quad (I.7)$$

### L'azimut solaire ( $\psi$ )

La position du soleil est généralement exprimée en termes d'azimut, qui est la projection vers le sud de la direction du soleil sur le plan horizontal est comptée comme ouest positif et est négatif [24], L'azimut du Soleil est donné par la relation [25].

$$\cos \psi = \frac{\sin h \sin \psi - \sin \delta}{\cos h \cos \psi} \quad (I.8)$$

### L'angle zénithal ( $\theta_z$ )

L'angle zénithal se définit comme l'angle engendré entre la trajectoire solaire et la direction du zénith de l'observateur, conformément à Szokolay[26]. Cette grandeur est exprimée par la relation suivante :

$$\theta_z = 90 - h \quad (I.9)$$

La Figure I.7 illustre la position du Soleil selon les coordonnées horizontales.

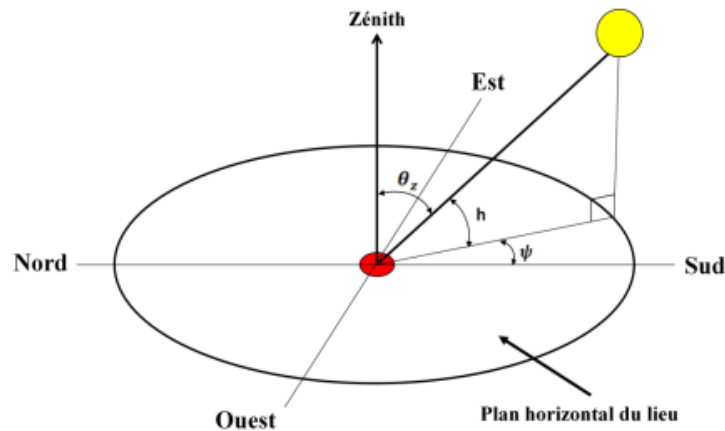


Figure I.7. Système de coordonnées solaires dans le plan horizontal [1].

## I.7. Impact de l'atmosphère sur l'irradiation solaire

### I.7.1. Atmosphère

L'atmosphère terrestre est constituée de divers composants, notamment 78 % d'azote, 21 % d'oxygène, 1 % d'argon, ainsi que d'autres gaz variés. À cela s'ajoutent la vapeur d'eau et les aérosols. Les aérosols sont des particules qui peuvent être d'origine naturelle, telles que le sable, la suie, les cristaux de glace ou les poussières volcaniques, ou d'origine anthropique, provenant par exemple de l'industrie, des véhicules automobiles ou du chauffage. L'atmosphère terrestre peut être subdivisée en plusieurs couches verticales en fonction des variations discontinues de température en relation avec l'altitude.

### I.7.2. Composition de l'atmosphère terrestre

L'atmosphère terrestre se compose de plusieurs couches superposées, chacune ayant des caractéristiques distinctes. De la surface de la Terre vers l'espace, ces couches sont :

- La troposphère (de 0 à environ 7-16 km) : C'est la couche où se déroulent les phénomènes météorologiques. La température diminue généralement avec l'altitude, avec une baisse d'environ  $-2^{\circ}\text{C}$  tous les 300 mètres en moyenne.
- La stratosphère (de 8-15 km à environ 50 km) : La température augmente progressivement jusqu'à environ 20 km, puis elle augmente plus rapidement pour atteindre un maximum d'environ 270 K vers 50 km d'altitude. Cette augmentation de température est due à l'absorption de l'irradiation solaire dans l'ultraviolet par l'ozone.



- La mésosphère (de 50 à 80 km) : La température diminue à mesure que l'on monte en altitude. La présence de dioxyde de carbone en quantité limitée provoque une baisse de la température jusqu'à atteindre des valeurs minimales d'environ 150 à 210 K.
- La thermosphère (de 80 jusqu'à environ 350-800 km) : Dans cette couche, l'atmosphère est chauffée par l'irradiation solaire ultraviolet de longueur d'onde inférieure à 175 nm.
- L'exosphère (de 350-800 km jusqu'à environ 50 000 km) : Cette couche représente la région extrêmement périphérique de l'atmosphère, située au-delà de la thermosphère, et s'étend dans l'espace.

Les diverses strates de l'atmosphère jouent un rôle fondamental dans le contrôle du climat et de l'impact de l'irradiation solaire sur la surface terrestre.

### I.7.3. Masse atmosphérique

L'intensité d'absorption et de diffusion subie par l'irradiation solaire lors de son passage à travers l'atmosphère est directement influencée par la densité de molécules qu'il rencontre. Cette relation est exprimée par la loi de transmission suivante :

$$Q = Q_0 \exp(-Kl) \quad (\text{I.10})$$

De telle sorte que :

$Q$  est le flux énergétique sortant après un parcours de longueur ( $l$ ) dans l'atmosphère

$Q_0$  est le flux énergétique incident

$K$  est le coefficient d'extinction caractérisant l'absorption et la diffusion

La masse atmosphérique (AM) représente l'épaisseur verticale moyenne de l'atmosphère au niveau de la mer, sous des conditions normales de température et de pression. La masse atmosphérique est calculée à l'aide de l'expression suivante :

$$MA = \frac{P}{1013} \frac{1}{\sin(h)} \exp\left(\frac{-Z}{7.8}\right) \quad (\text{I.11})$$

Où

$P$  est la pression de l'air ( $\text{N/m}^2$ )

$h$  est la hauteur solaire, sur l'horizon ( $h = 90^\circ$  au zénith)

$Z$  est l'altitude en (km)

Dans les conditions normales ( $p = 1013 \text{ N/m}^2$ ) et au niveau de la mer ( $Z = 0$ ), l'expression de la masse atmosphérique est simplifiée à :

$$AM = \frac{1}{\sin(h)} \quad (\text{I.12})$$

### I.7.4. Indice de clarté

Le rapport entre l'irradiation au sol et l'irradiation extraterrestre est nommé indice de clarté. L'indice de clarté,  $K_T$ , est défini par [27]:

$$K_T = \frac{G}{G_0} \quad (\text{I.13})$$

Où :  $G$  est l'irradiation solaire reçue sur un plan horizontal

La valeur moyenne mensuelle de l'indice de clarté  $K_T$  présente des variations en fonction de la localisation géographique et de la saison, oscillant habituellement entre 0,3 (dans les régions ou les périodes pluvieuses) et 0,8 (dans les régions ou les périodes sèches et ensoleillées).

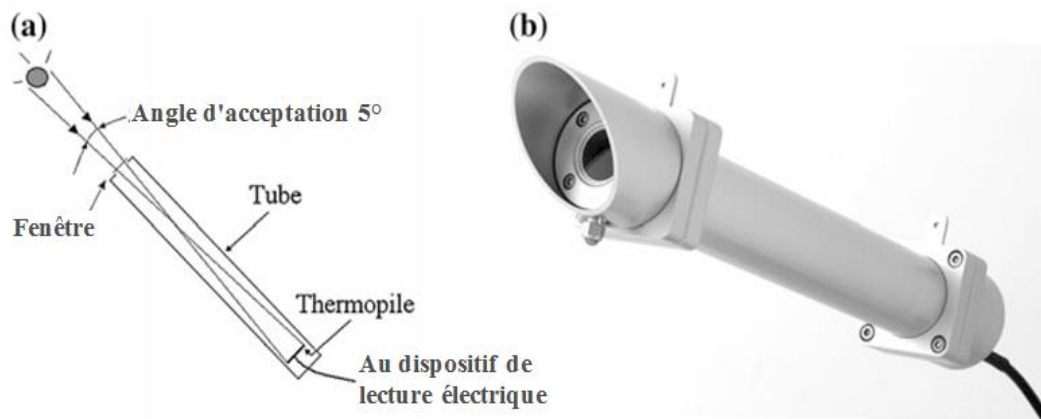
### I.8. Mesure d'irradiation solaire

Cette section est conçue comme une introduction concise pour l'instrumentation utilisée pour mesurer les composants du rayonnement solaire ainsi que pour les caractéristiques climatiques et la localisation géographique des zones où les données observées ont été enregistrées.

#### I.8.1. Radiomètres solaires

##### Pyrhéliomètre

Il est conçu pour mesurer l'irradiation solaire directe à incidence normale. Il se compose d'un thermophile enroulé avec un fil noir au fond d'un tube avec une ouverture (généralement entre  $5^\circ$  et  $6^\circ$  angle plein total) dans un rapport de 1 à 10 sur sa longueur, monté sur un système de repérage automatique contrôlé par un ordinateur. L'irradiation diffusée est bloquée par la surface du détecteur. Lorsque le soleil est obscurci par les nuages, cette magnitude est nulle [24] est présentée à la figure. I.8

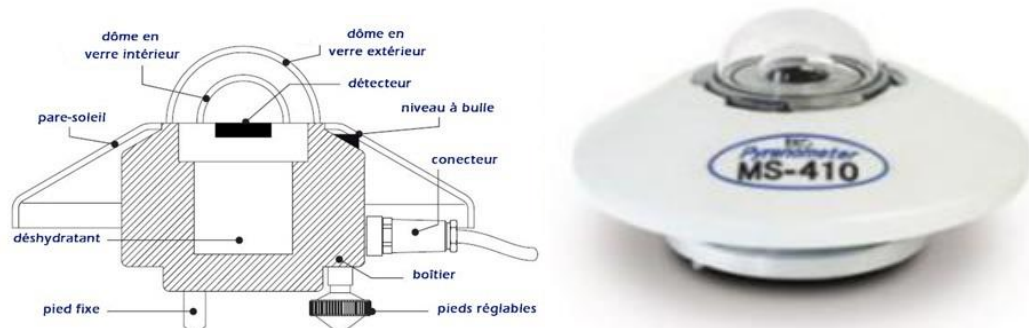


**Figure. I.8.** a) Schéma d'un pyrhéliomètre. b) Photo de Hukseflux DR01 [28].

### Pyranomètre

Il mesure l'irradiation solaire globale ou diffuse généralement sur une surface horizontale. Il se compose d'un pyranomètre équipé d'une série de thermocouples, alternant entre des soudures chaudes et des soudures froides. Lorsque l'instrument est exposé au rayonnement solaire, il crée une différence de température. Cette différence est détectée par le pyranomètre, qui convertit l'énergie thermique en énergie électrique. La tension de sortie peut ensuite être mesurée et convertie en valeurs de rayonnement solaire ( $W/m^2$ ).

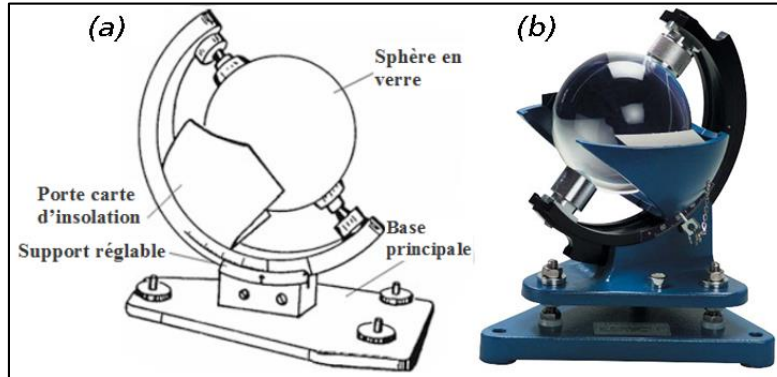
Pour mesurer spécifiquement l'irradiation diffuse, un pare-soleil sous forme d'une bande métallique circulaire est ajouté pour protéger le pyranomètre de l'irradiation solaire direct. Cela permet de mesurer uniquement l'irradiation solaire diffus qui provient de toutes les directions autres que le soleil lui-même (Figure I.9) [29].



**Figure I.9.** a) Schéma d'un pyranomètre. b) Pyranomètre type EKO MS-410 [15].

### Mesure de la durée d'ensoleillement

La durée d'ensoleillement d'une journée se définit comme la période pendant laquelle l'éclairement solaire direct atteint ou dépasse un seuil conventionnel de  $120 Wm^2$ . En l'absence de nuages permanents, la durée d'ensoleillement est généralement similaire à la durée du jour. Pour obtenir les durées d'ensoleillement mensuelles et annuelles, on cumule simplement les durées d'ensoleillement de chaque journée. L'instrument utilisé pour mesurer la durée d'ensoleillement est appelé un héliographe (Figure. I.10). Cet appareil enregistre les périodes où l'intensité de l'irradiation solaire direct dépasse le seuil mentionné, fournissant des lectures en heures et en plus.



**Figure I.10.** a) Schéma d'héliographe Campbell-Stokes. b) Photo typique d'héliographe Campbell-Stokes) [28].

### Albédomètre

L'albédo est formé par deux pyranomètres identiques disposés en opposition. Le premier pyranomètre est orienté vers le haut pour mesurer l'irradiation global, tandis que le second est orienté vers le bas pour mesurer l'irradiation réfléchi par le sol (figure I.11).



**Figure I.11.** Capteur Albédomètre [1] .

## I.9. Mesures Satellites météorologiques

Contrairement aux mesures locales au sol, les mesures par satellite couvrent généralement de vastes zones, de sorte que des mesures précises du rayonnement solaire à plusieurs endroits peuvent être obtenues à l'aide du même instrument. En effet, cette technique permet de suivre l'évolution spatiale et temporelle des nuages et d'étudier les bilans énergétiques terrestres. Actuellement, les satellites géostationnaires collectent des données toutes les 15 minutes et ont une résolution spatiale faible à moyenne (environ 1 km), tandis que les satellites en orbite offrent une résolution spatiale plus élevée, Les satellites météorologiques sont disponibles en trois configurations orbitales principales : équatoriaux, polaires et géostationnaires. Les satellites équatoriaux sont des satellites en orbite basse (~ 600 km

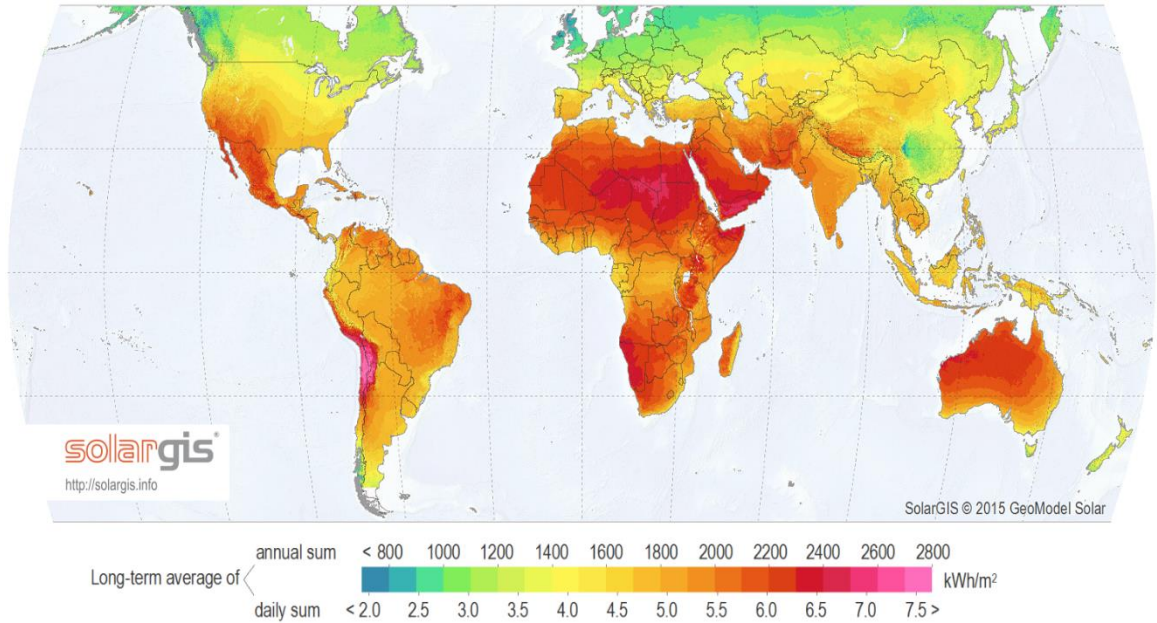
d'altitude) qui tournent généralement en orbite autour de la terre dans une direction sinusoïdale qui traverse l'équateur au moins deux fois par orbite. Les satellites polaires sont également des satellites en orbite basse qui orbitent autour de la terre du pôle nord au pôle sud. Les satellites géostationnaires sont mis en orbite de telle manière qu'ils sont toujours au même point sur la surface de la terre. Les satellites géostationnaires sont très hauts (environ 36 000 km) et peuvent fournir des images à haute résolution temporelle de vastes zones de la surface de la terre.

Pour établir des mesures à distance, les satellites météo utilisent des capteurs passifs (comme les radiomètres), qui fixent simplement le milieu en collectant les ondes qu'il émet : c'est la télédétection passive [30], ou il peut utiliser des capteurs actifs comme les radars (lidars, interféromètre, diffusiomètre). Les capteurs actifs génèrent leur propre énergie pour éclairer la cible. Le capteur émet un rayonnement sur l'objet observé, puis capte et mesure le rayonnement réfléchi par l'objet.

Il offre donc l'avantage de pouvoir prendre des mesures à n'importe quel moment de la journée ou de la saison. Cependant, ils doivent cependant produire une quantité considérable d'énergie pour bien éclairer une cible : c'est la télédétection active [27]. Un satellite n'emporte pas qu'un radiomètre mais dispose, le plus souvent de plusieurs instruments de mesure qui opèrent dans différents domaines (différentes bandes spectrales, différentes résolutions spatiale et temporelle...). Plusieurs méthodes d'estimation du rayonnement solaire à partir de données satellitaires ont été améliorées au cours des dernières décennies, la plupart utilisant l'imagerie satellitaire géostationnaire [31].

### **I.10. Rayonnement solaire dans le monde**

Figure I.12 montre la distribution du rayonnement solaire dans le monde. L'énergie annuelle qui peut être captée en  $1\text{m}^2$  est exprimée par l'irradiation normale directe. La quantité d'énergie solaire qui brille à chaque endroit est différente. En raison des variations climatiques entre les différentes régions du monde, il est attendu que des différences significatives se manifestent en ce qui concerne l'irradiation solaire, il peut être divisé en trois régions [32].



**Figure I.12.** Intensité de l'irradiation solaire mondial [33].

- La première région, située entre les latitudes 15 N et 35 N, bénéficie de la plus grande quantité de l'irradiation solaire. Cette abondance de l'irradiation solaire est principalement attribuable à l'irradiation directe, en raison de la faible couverture nuageuse et de la rareté des précipitations dans cette zone.
- La deuxième région, qui s'étend de 15 N à 0 N, présente le deuxième niveau le plus élevé de l'irradiation solaire. Cette région se caractérise par des niveaux plus élevés de l'irradiation diffus en raison de l'humidité et de la présence de nuages. Dans cette zone, l'irradiation solaire reste relativement constante tout au long de l'année.
- La troisième région, située au-dessus de 35 N, enregistre le niveau le plus bas de l'irradiation solaire. Cela est principalement dû à une couverture nuageuse plus dense. Dans cette région, près de la moitié de l'irradiation provient de l'irradiation diffuse. l'irradiation solaire est également plus faible aux pôles de la Terre.

Comme le montre cette figure, une région du Sahara de cette taille pourrait alimenter le monde entier grâce à l'énergie solaire. Au fil des ans, différents projets ont été proposés pour faire de cette idée une réalité, mais il est toujours difficile de stocker et de transporter autant d'électricité à partir d'un endroit aussi éloigné. En théorie, une petite partie du désert pourrait effectivement alimenter une grande partie du monde. Mais comment cela serait-il réalisé ? Aujourd'hui, deux technologies dominent l'industrie de l'énergie solaire : l'énergie solaire concentrée (CSP) et l'énergie photovoltaïque (PV).

- Les systèmes à énergie solaire concentrée (CSP) produisent de l'énergie électrique en convertissant l'énergie solaire en chaleur à haute température à l'aide de réflecteurs pour concentrer l'énergie solaire en un seul endroit afin de conduire un moteur thermique, puis convertit cette énergie thermique en électricité.
- Les panneaux solaires photovoltaïques (PV) sont beaucoup plus flexibles et plus faciles à installer. Les panneaux solaires pourraient avoir un impact remarquable sur le désert. Contrairement au CSP qui utilise l'énergie solaire, les panneaux solaires photovoltaïques utilisent plutôt la lumière du soleil. En d'autres termes, le photovoltaïque est la conversion directe de la lumière en électricité. Les panneaux solaires photovoltaïques pourraient-ils affecter le sahara? Des études de modélisation antérieures ont montré que les fermes solaires à grande échelle peuvent produire des changements climatiques importants à l'échelle continentale. Lorsque les panneaux solaires réfléchissent moins de chaleur dans l'espace que le sable. Cela signifie que la surface se réchaufferait, causant l'air pour se lever et former des nuages. Cela signifierait plus de précipitations. Bien que l'installation massive de panneaux solaires au Sahara pourrait également avoir un impact remarquable sur le désert lui-même, mais moins efficace par temps très chaud [34].

### I.10.1 Potentiel de l'énergie solaire en Algérie

Compte tenu de sa situation géographique, l'Algérie possède l'un des meilleurs gisements solaires au monde. Presque tous les territoires nationaux ont plus de 3000 heures d'ensoleillement par an et peuvent atteindre 3900 heures dans les hautes terres et le désert du sahara. L'énergie quotidienne totale annuelle obtenue au niveau de  $1 \text{ m}^2$  dans la plupart des régions du pays dépasse 5000 Wh par an, et elle dépasse 5700 Wh dans la partie sud du pays (Figure. I.13). La production totale d'électricité dans le nord est comprise entre 1800 et 1900 kWh/m<sup>2</sup>/an, et dans le sud du pays entre 2000 et 2200 kWh/m<sup>2</sup>/an [35].

La répartition du potentiel solaire selon les différentes régions climatiques en Algérie est illustrée dans la figure I.14 sur la base d'évaluations satellitaires, l'Agence spatiale allemande (ASA) a conclu que l'Algérie représente le plus grand potentiel d'énergie solaire de tout le bassin méditerranéen, à savoir : 169 000 TWh/an de l'énergie thermique, 13,9 TWh/an de l'énergie photovoltaïque, 35 TWh/an de vent. La distribution du potentiel solaire dans différentes régions climatiques est illustrée à la figure I.13.

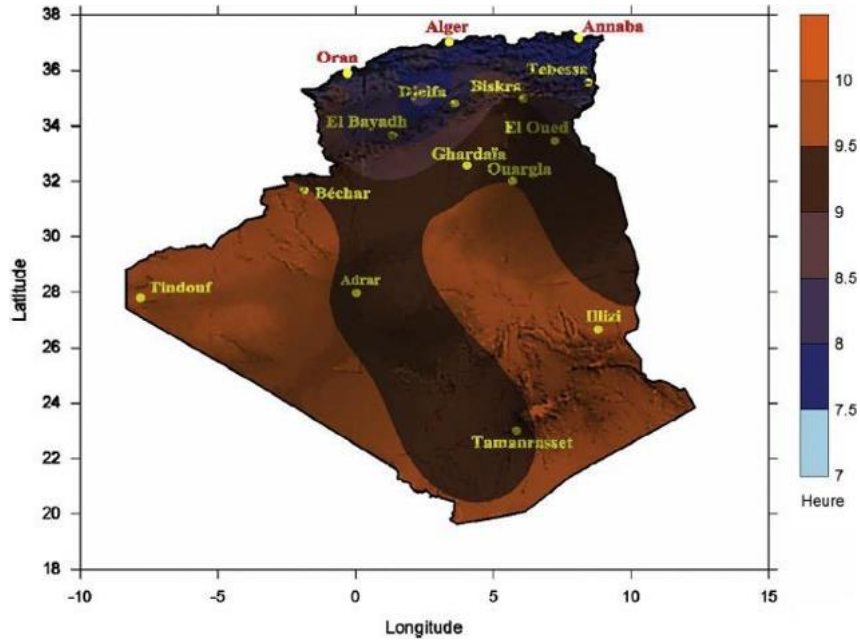


Figure I.13. Durée d'insolation en Algérie [36].

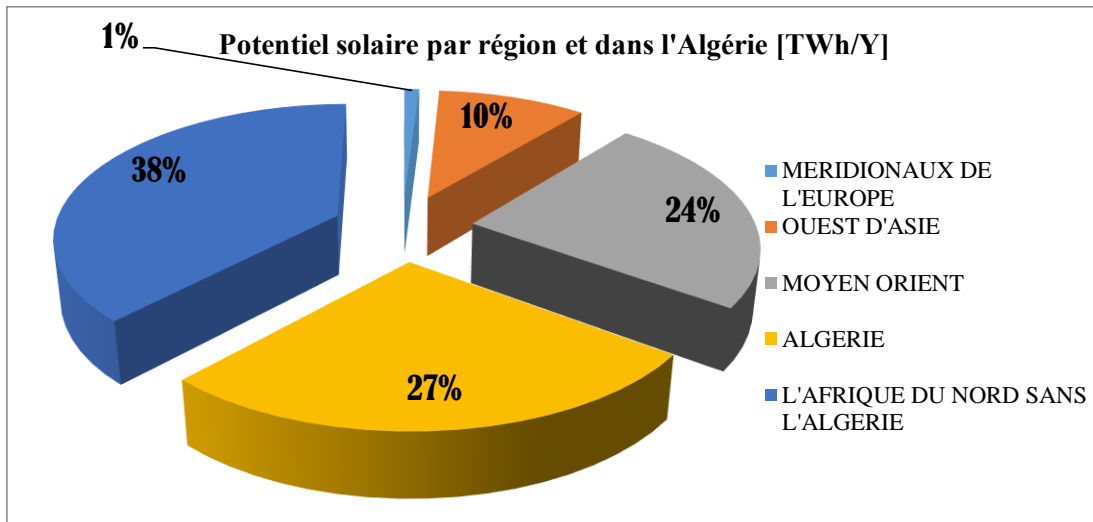


Figure I. 14. Présentation des ratios du potentiel solaire [19].

Le tableau I.1 présente la distribution du potentiel solaire par région climatique sur le territoire algérien, en fonction de l'ensoleillement annuel reçu.



**Tableau I.1.** Distribution de l'énergie solaire en Algérie [37].

Régions	Région côtière	Hauts Plateaux	Sahara
Surface (%)	4	10	86
Durée d'insolation (heures/an)	2650	3000	3500
Energie reçue (kWh/m <sup>2</sup> /an)	1700	1900	2650

### **I.10.2. Programme de développement des énergies renouvelables en Algérie**

L'Algérie entame une nouvelle ère d'énergie durable. Le plan d'énergie renouvelable mis à jour comprend l'installation d'environ 22 000 mégawatts d'énergie renouvelable pour le marché national d'ici 2030, avec le maintien des options d'exportation comme objectif stratégique si les conditions du marché le permettent. À travers ce plan énergies renouvelables, l'Algérie entend se positionner comme un acteur majeur de la production photovoltaïque et éolienne en intégrant la biomasse, la cogénération, la géothermie et le solaire thermique au-delà de 2021. Ces filières énergétiques seront des moteurs de développement économique durable capables de porter de nouveaux modèles de croissance économique. D'ici 2030, 37 % de la capacité installée et 27 % de la consommation nationale d'électricité proviendront de sources d'énergie renouvelables.

Selon le site du ministère de l'Énergie, le potentiel national d'énergies renouvelables est dominé par l'énergie solaire, que l'Algérie considère comme une opportunité et un levier de développement économique, notamment à travers la création d'industries créatrices de richesses et d'emplois. La stratégie de l'Algérie dans ce domaine vise à développer une véritable industrie des énergies renouvelables liée à des programmes de formation et de capitalisation des connaissances, ce qui permettra à terme d'employer le génie local algérien, notamment dans l'ingénierie et la gestion de projet. Le programme des énergies renouvelables (EnR) cible la demande d'électricité sur le marché national et créera des milliers d'emplois directs et indirects. La consistance des plans d'énergies renouvelables mis en œuvre pour le marché national sur la période 2015-2030 est de 22 000 MW, le tableau I.2 présente la répartition par filière.

**Tableau I.2.** Consistance du programme de développement des EnR, [23].

	2015-2020 (MW)	2021-2030 (MW)	Total [MW]
<b>Photovoltaïque</b>	3000	10575	13575
<b>Eolien</b>	1010	4000	5010
<b>CSP (solaire thermique)</b>	---	2000	2000
<b>Cogénération</b>	150	250	400
<b>Biomasse</b>	360	640	1000
<b>Géothermie</b>	05	10	15
<b>Total</b>	4525	17475	22000

En ce qui concerne les systèmes énergétiques connectés au réseau, le soutien gouvernemental est mis en œuvre au moyen du mécanisme de tarif d'achat garanti: Ce dispositif assure aux producteurs d'énergie renouvelable la certitude de recevoir des tarifs qui garantissent une rentabilité raisonnable pour leur investissement sur une période d'éligibilité de 20 ans dans le cas du solaire et de l'éolien, et de 15 ans pour la cogénération.

Afin d'atteindre les objectifs fixés dans ce programme, tous les moyens pédagogiques et matériels doivent être mis à la disposition des chercheurs dans le domaine des énergies renouvelables. La plus importante d'entre elles est peut-être de fournir aux chercheurs une base de données sur le climat. Malheureusement, le réseau radiométrique du pays n'est pas très dense par rapport à la taille du territoire. En effet, sur les 60 stations météo qui composent le réseau l'office national de la météorologie (ONM), seules 7 sont garanties pour mesurer les composantes diffuse et globale de l'irradiation solaire reçue à l'horizontale, et très peu sont garanties pour enregistrer la composante directe. Pour cela, nous avons choisi d'étudier au cours de cette thèse, autres méthodes de calcul des prédictions de données météorologique requises [24].

### **I.11. Conclusion**

Dans ce chapitre nous avons rappelé les bases théoriques de la caractérisation des irradiances solaires comme les coordonnées horaires et terrestres, la principale source d'alimentation de tous les systèmes de conversion de l'énergie solaire est l'irradiation solaire, depuis son émission par le soleil jusqu'à son arrivée à la surface terrestre. Nous avons également présenté les différentes principales instrumentations de mesures de l'irradiation solaire utilisées dans les centres de recherche des énergies renouvelables et des industries solaires.

Le prochain chapitre abordera la présentation de différentes méthodes utilisées pour la prédiction de l'irradiation solaire reçue à la surface terrestre.

# Chapitre II

Synthèse bibliographique sur la modélisation de l'irradiation solaire  
par les modèles d'intelligence artificielle

## **II.1. Introduction**

La revue de la littérature joue un rôle fondamental dans la recherche scientifique en permettant aux chercheurs d'utiliser les connaissances passées comme point de départ pour leurs propres études. Cependant, il est essentiel de ne pas entreprendre cette démarche de manière isolée, mais plutôt de l'intégrer dans un contexte plus large qui prend en compte les problèmes liés au sujet d'étude. Cette approche vise à fusionner les idées préexistantes avec les connaissances actuelles pour obtenir des résultats innovants.

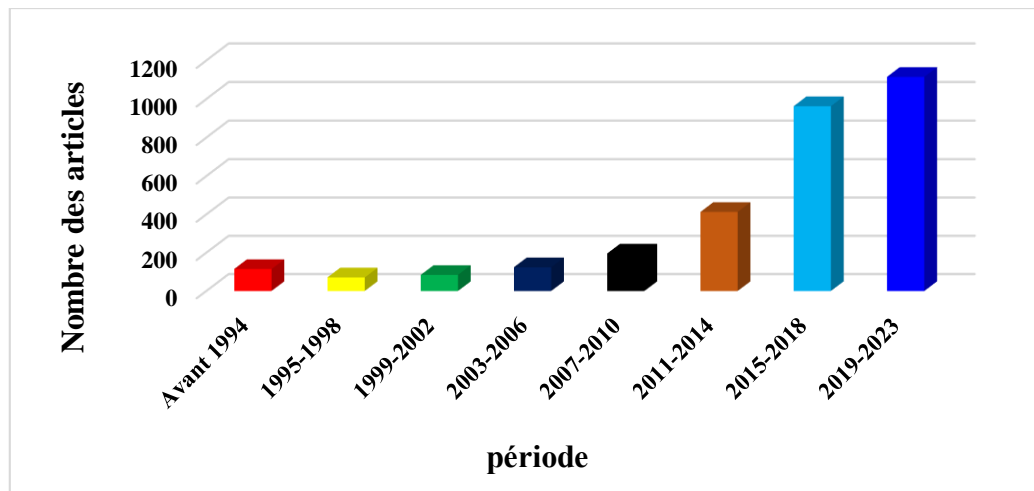
L'indisponibilité et la qualité parfois insuffisante des données sur l'irradiation solaire dans certaines stations météorologiques, en particulier dans les pays en développement, sont des obstacles à la recherche. La modélisation de l'irradiation solaire est donc essentielle, et de nombreuses approches sont utilisées pour estimer ou prédire l'irradiation solaire en se basant sur des variables météorologiques plus accessibles.

Ce chapitre se penche sur cette problématique en examinant la littérature existante, mettant en évidence les diverses approches de modélisation de l'irradiation solaire. Il montre également qu'il existe plusieurs façons d'aborder ce sujet, offrant ainsi une variété d'options aux chercheurs.

## **II.2. Production scientifique mondiale sur la prédiction de l'irradiation solaire**

Nous avons commencé par mener une étude bibliométrique sur les publications scientifiques axées sur l'irradiation solaire avant d'explorer plus en détail les diverses méthodes de modélisation. Cette analyse, outre son importance pour l'histoire de la science, nous a permis d'évaluer la croissance, la qualité, la quantité et les domaines d'intérêt des chercheurs dans ce domaine.

La Figure II.1 synthétise l'évolution du nombre d'articles publiés dans ce domaine au fil du temps en utilisant "ISI Web of Knowledge" et "Google Scholar" comme sources de référence. Une croissance remarquable de la production scientifique est observable au cours des dernières années. Notre méthodologie repose sur l'exploration des mots-clés présents dans les publications antérieures, ce qui nous a amenés à élaborer une liste de mots-clés pour chaque technique de modélisation, tout en évitant la duplication des publications par une approche logique appropriée.



**Figure II.1.** Nombre de publications inhérents à la prédiction d'irradiation solaire publiés durant les années avant 1994 à 2023 (consulté juin 2023).

### II.3. Intelligence artificielle

L'intelligence artificielle (IA) est une discipline avancée de l'informatique qui vise à créer des automates capables de fonctionner de manière similaire à la pensée humaine. Elle englobe toutes les modélisations de connaissances et les techniques visant à atteindre cet objectif. Les techniques d'IA deviennent de plus en plus populaires en tant qu'approche alternative aux méthodes conventionnelles ou en tant que composants de systèmes intégrés [39]. Elles ont été utilisées pour résoudre des problèmes complexes dans divers domaines pratiques. Voici quelques-uns des domaines d'application de l'IA (Vision par ordinateur, Santé et médecine, Industrie manufacturière, Finance, Transport et logistique, Commerce électronique, Énergie et environnement, Robotique, Jeux, Éducation, Agriculture...etc.) [40]. Ces succès sont dus particulièrement à la multiplication des jeux de données, le développement de la puissance de calcul des ordinateurs et l'amélioration des algorithmes.

L'intelligence artificielle (IA) est composée de plusieurs branches, telles que les machines à vecteurs de support (MVS), les réseaux de neurones artificiels (RNA), les algorithmes génétiques (GA), ainsi que divers systèmes hybrides (HS) qui combinent deux ou plusieurs de ces branches mentionnées précédemment. [41]. Enfin, nous nous intéressons à deux techniques bien connues dans la modélisation de l'irradiation solaire qui sont :

- Réseaux de neurones artificiels (RNA)
- Machine à vecteur de support (MVS)

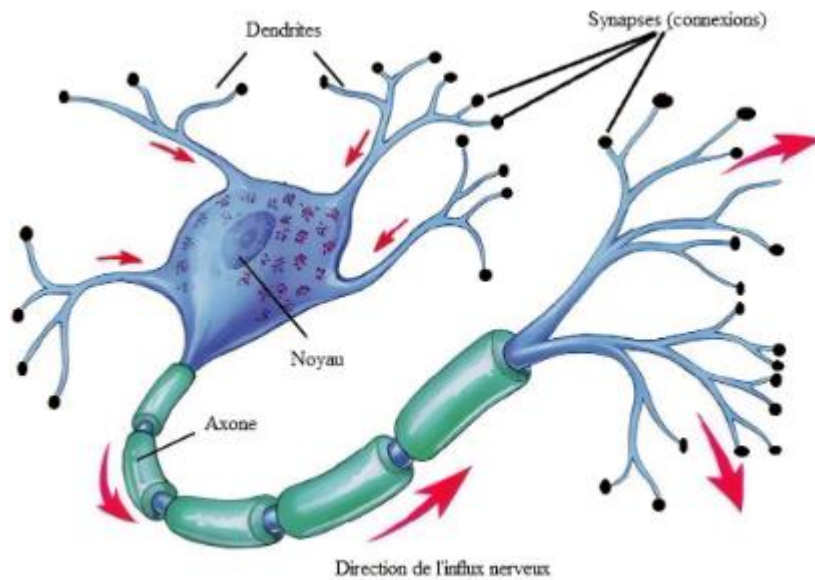
### **II.3.1. Réseaux de neurones artificiels (RNA)**

#### **II.3.1.1. Neurone biologique**

Le neurone considéré comme l'unité fondamentale du système nerveux central, se compose d'un corps cellulaire appelé soma, à partir duquel des ramifications se forment, ce que l'on appelle les dendrites (figure II.2).

Les dendrites jouent le rôle de transporteurs d'informations provenant de l'extérieur ou d'autres neurones, et les transmettent ensuite vers le soma. Une fois que le soma a traité cette information, elle est transmise aux autres neurones à travers un prolongement de sortie nommé axone.

La communication entre un axone et une dendrite se réalise au travers d'un espace vide connu sous le nom de synapse, où la transmission s'effectue par l'intermédiaire de substances chimiques.



**Figure II.2.** Représentation d'un neurone biologique [42].

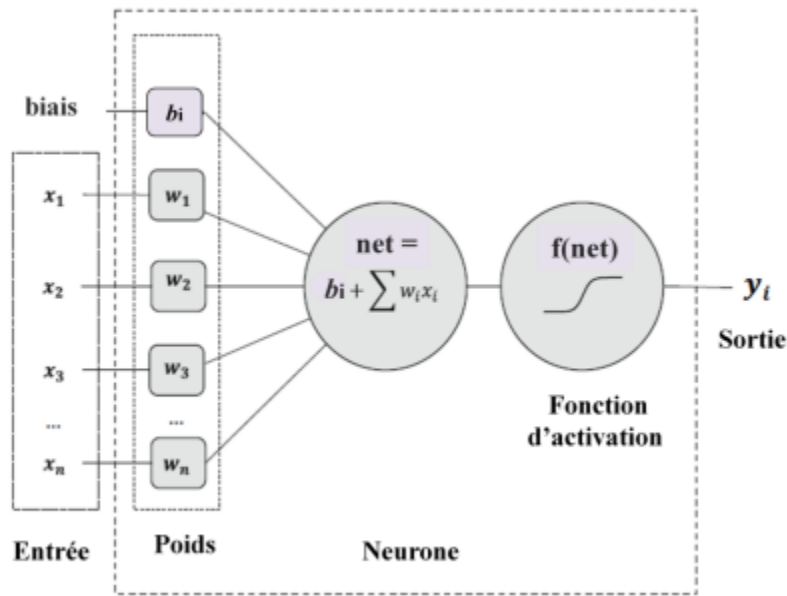
#### **II.3.1.2. Neurone artificiel (formel)**

Le neurone artificiel est un concept de calcul qui tire son inspiration du fonctionnement des neurones biologiques. Il s'agit d'un élément formel agissant comme un opérateur, recevant diverses entrées provenant du milieu extérieur ou d'autres neurones. Chacune de ces entrées est multipliée par un poids synaptique spécifique, puis elles sont sommées pour produire une sortie. Cette sortie n'est générée que lorsque la somme pondérée des entrées dépasse un seuil interne préétabli.

Pour évaluer la sortie du neurone, une somme pondérée des entrées est généralement calculée, puis le résultat est soumis à une fonction non linéaire. Mathématiquement, ce processus peut être décrit par les équations suivantes :

$$y = w_0 + \sum_{i=1}^n (w_i x_i) \quad (\text{II.1})$$

Notons que le poids  $w_0$  est attribué à une entrée constante connue sous le nom de "biais". Ce potentiel est ensuite soumis à une fonction  $f$ , appelée fonction d'activation. Le choix de la fonction d'activation peut être crucial dans la construction des réseaux de neurones. Pour représenter conceptuellement le fonctionnement du neurone biologique, on utilise le modèle illustré dans la figure II.3. Contrairement aux neurones biologiques dont les états sont binaires, la majorité des fonctions de transfert sont continués, ce qui permet d'obtenir une infinité de valeurs possibles dans l'intervalle  $[0, +1]$  ou  $[-1, +1]$ .



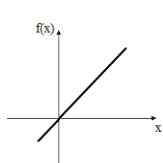
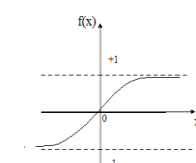
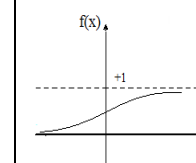
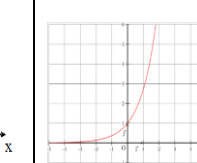
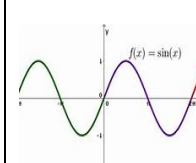
**Figure II.3.** Représentation mathématique d'un neurone artificiel [1].



**II.3.1.3. Fonctions d'activation**

La fonction d'activation joue un rôle essentiel en restreignant l'amplitude du signal de sortie du neurone, imitant ainsi l'effet de seuil caractéristique des neurones biologiques. Au fil du temps, les fonctions de transfert ont évolué et différentes fonctions mathématiques sont couramment utilisées dans nos recherches, comme indiqué dans le tableau II.1.

**Tableau II.1.** Fonctions de transfert usuelles,  $f(x)$  [43].

Fonction de transfert					
	Linéaire	Logistique (Sigmoidale logarithmique)	Tangente hyperbolique	Exponentielle	sinus
<b>Notation</b>	Purelin	Logsig	Tansig	Exponential	sin
<b>Formule de fonction de transfert</b>	$f(x) = x$	$f(x) = \frac{1}{1 + e^{-x}}$	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$f(x) = e^x$	$f(x) = \sin(x)$
<b>L'allure de la fonction</b>					

**II.3.1.4. Architecture d'un réseau de neurones**

Un réseau de neurones (RNA) est constitué d'un ensemble de neurones artificiels interconnectés, et on peut distinguer principalement deux types d'architectures [44]:

- Réseaux de neurones non bouclés : ces réseaux effectuent une ou plusieurs fonctions algébriques sur les entrées en combinant les fonctions exécutées par chaque neurone. Ils servent principalement à approximer des fonctions non linéaires et à modéliser des processus statiques non linéaires.
- Réseaux de neurones bouclés : ces réseaux constituent des systèmes dynamiques "à temps discret" régis par une ou plusieurs équations aux différences non linéaires, impliquant la fonction réalisée par chaque neurone et le délai associé à chaque connexion. Ils sont utilisés pour la modélisation,

énergie, prévision météorologique, robotique, industrie, Sciences de l'environnement etc...

Dans notre étude, nous avons choisi d'utiliser le modèle le plus couramment utilisé, le perceptron multicouche (PMC), comme illustré dans la figure II.4. Les principes de fonctionnement de ce modèle seront expliqués dans les sections suivantes.

### II.3.1.5. Perceptron multicouches

Le perceptron monocouche (PMC) est un réseau de neurones non bouclé avec une structure très spécifique. Ces neurones sont organisés en couches successives, où l'information circule dans un sens, de la couche d'entrée vers la couche de sortie sans retour en arrière. Les neurones d'une même couche ne sont pas connectés les uns aux autres. Un neurone ne peut envoyer ses résultats qu'aux neurones d'une couche après lui. Le réseau est constitué de plusieurs couches. La figure II.4 illustre la structure de ce type de réseau.

- Couche d'entrée: Cette couche est associée aux données d'entrée du système et est composée de cellules d'entrée qui représentent les  $i$  variables d'entrée.
- Couche(s) cachée (s): Chaque couche cachée est constituée de  $j$  neurones. Ces neurones n'ont pas de lien direct avec l'extérieur et sont appelés neurones cachés, car ils ne sont pas directement liés aux entrées ou aux sorties du système.
- Couche de sortie: Cette couche est composée des valeurs de sortie résultant de la combinaison linéaire des fonctions de transfert des neurones de la couche cachée. C'est la couche qui donne les résultats finaux du réseau en fonction des données d'entrée et des poids appris lors de l'apprentissage du réseau.

Les neurones sont interconnectés au moyen de connexions pondérées. Ce sont les valeurs de ces poids de connexion qui régissent le fonctionnement du réseau, lui permettant d'effectuer une transformation non linéaire de l'espace des entrées vers l'espace des sorties [44].

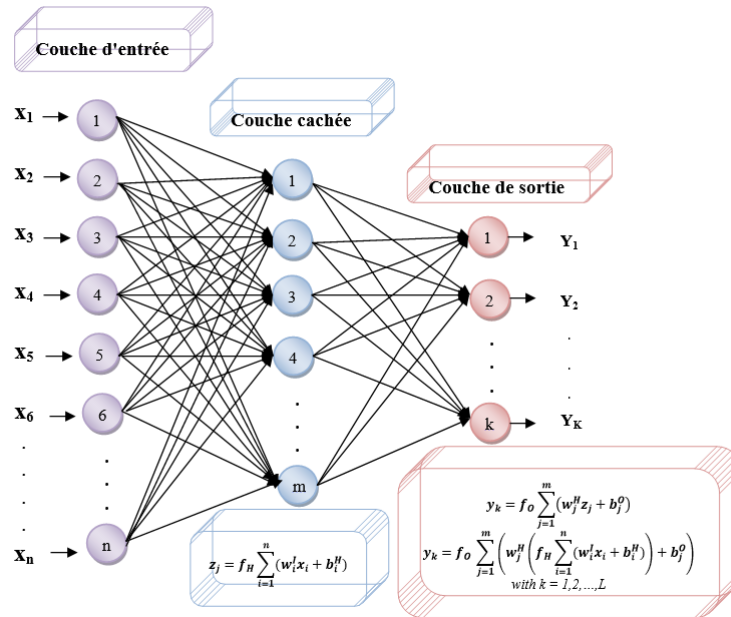


Figure II.4. Réseau de neurones non bouclé.

### II.3.1.6. Apprentissage

L'apprentissage est le processus par lequel les valeurs des poids synaptiques du réseau sont déterminées, et sa structure est établie. Cela se réalise au moyen d'un algorithme d'apprentissage, qui est un type spécifique d'algorithme d'optimisation, tels que la programmation non-linéaire, la programmation génétique, l'analyse d'intervalle ou le recuit simulé, en utilisant également des méthodes de substitution des différences finies pour les dérivées [43].

Plusieurs algorithmes d'apprentissage implémentent la rétro propagation. Cependant, chaque algorithme présente ses propres avantages, tels que le taux de calcul et les exigences de calcul et de mémoire, mais il n'existe pas d'algorithme universellement adapté à tous les problèmes. La performance de chaque algorithme dépend du processus modélisé, de l'échantillon d'apprentissage et du mode d'apprentissage. Par conséquent, le choix des algorithmes est d'une importance primordiale et étroitement lié à l'application spécifique, reposant principalement sur des expérimentations [43].

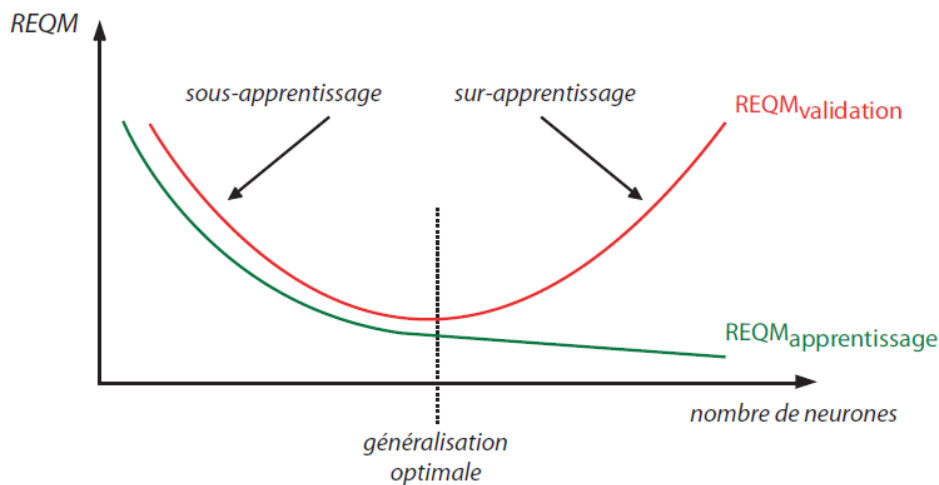
L'algorithme de rétro propagation du gradient se compose de deux phases :

- Phase de propagation: Cette étape implique la présentation d'une configuration d'entrée au réseau, suivie de sa transmission à travers les différentes couches cachées jusqu'à la sortie.

- Phase de rétro-propagation : Après la phase de propagation, l'objectif consiste à réduire au maximum l'erreur globale observée sur l'ensemble des exemples présentés. Cette erreur est déterminée comme une fonction des poids synaptiques et équivaut à la somme des carrés des différences entre les sorties calculées par le réseau et les valeurs cibles pour tous les exemples contenus dans le jeu de données d'apprentissage.

### **II.3.1.7. sur-apprentissage et sous-apprentissage**

Un bon modèle d'intelligence artificiel nécessite un compromis entre le biais et la variance du modèle. La figure II.5 représente la variation de l'erreur de prédiction en fonction de la complexité du modèle. Dans le cas des méthodes (AI) avec une variance élevée (méthodes complexes), le modèle est trop ajusté aux données. Dans ce cas, le modèle apprend même le bruit présent dans les données d'apprentissage et cela entraîne des conséquences négatives sur la capacité prédictive du modèle face à de nouvelles données (**sur-apprentissage**). En revanche, des algorithmes avec un biais élevé produisent des modèles plus simples et plus pauvres en termes de qualité de prédiction (**sous-apprentissage**)

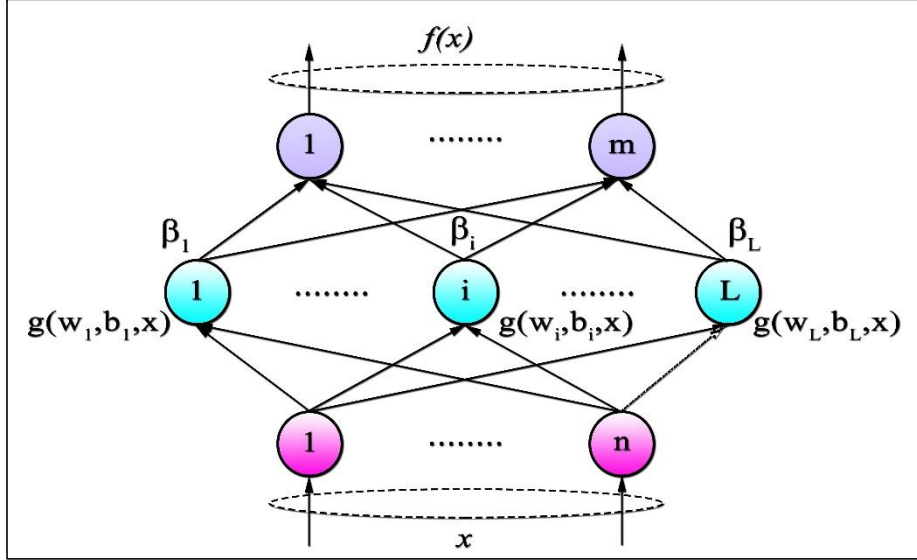


**Figure II.5.** Compromis biais-variance d'un modèle optimisé.

### **II.3.1.8. Machine D'apprentissage Extrême (ELM)**

La machine d'apprentissage extrême (Extreme Learning Machine, ELM) est un réseau neuronal non récurrent à une seule couche cachée proposé par Huang et al [45]. Dans un modèle ELM, les paramètres de la couche cachée sont générés aléatoirement et ne nécessitent pas d'être ajustés alors que les poids de

sorties sont déterminés analytiquement. Ceci a pour avantage de diminuer fortement le temps d'apprentissage [17]. La figure II.6 montre la structure d'un modèle ELM.



**Figure II.6.** Structure d'un modèle ELM [17].

Considérons un ensemble de données  $\{x_i, y_i\} \in R^n \times R^m \quad i = 1: N$ . La sortie  $y_i$  du modèle ELM est donnée par :

$$y_i = \sum_{j=1}^L \beta_j g(w_j x_i + b_j) \quad (\text{II.2})$$

où  $w_j = [w_{j1}, w_{j2}, \dots, w_{jn}]$  représente le vecteur de poids qui relie le  $j^{\text{ème}}$  nœud caché aux nœuds d'entrée,  $\beta_j = [\beta_{j1}, \beta_{j2}, \dots, \beta_{jm}]$  est le vecteur de poids qui relie les nœuds de la couche cachée aux nœuds de sortie,  $b_j$  est le biais du  $j^{\text{ème}}$  nœud caché et  $g$  est la fonction d'activation qui peut être linéaire, sigmoïde, gaussienne ou tangente hyperbolique.

L'équation peut être écrite sous la forme :

$$H\beta = y \quad (\text{II.3})$$

avec  $H$  la matrice de sortie de la couche cachée du modèle ELM donnée par :

$$H = \begin{bmatrix} g(w_1 x_1 + b_1) & \cdots & g(w_L x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(w_1 x_N + b_1) & \cdots & g(w_L x_N + b_L) \end{bmatrix} \quad (\text{II.4})$$

En conclusion, l'algorithme ELM peut être résumé en trois étapes. D'abord, les poids  $w_j$  et les biais  $b_j$  sont générés aléatoirement. Ensuite, la matrice  $H$  est calculée. Enfin, les poids de sorties  $\beta$  sont déterminés comme suit :

$$\beta = H^+ y \tag{II.5}$$

Avec  $H^+$  représente l'inverse généralisé de Moore-Penrose de la matrice  $H$  donné par :

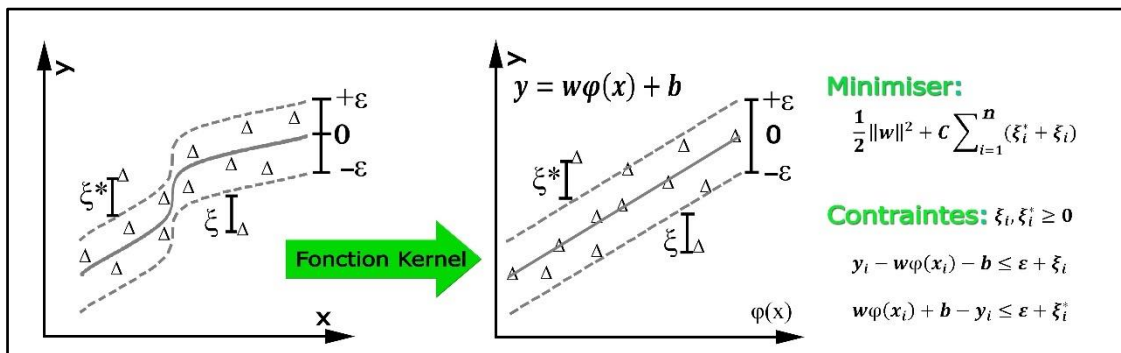
$$H^+ = (H^T H)^{-1} H^T \quad . \tag{II.6}$$

Les principaux hyperparamètres de ce modèle sont (activation\_func) : la fonction d'activation et (n\_hidden) : le nombre de neurones dans la couche cachée.

### II.3.2. Machines à vecteurs support (MVS)

Les machines à vecteurs support (MVS) ont été proposées par Vapnik en 1995 pour résoudre des tâches de classification. Elles ont été ensuite étendues pour la régression, elles sont nommées dans ce cas : Machines à vecteurs support pour la régression (MVSR). Les MVSR s'appuient sur la notion de réduction du risque structurel, telle que définie dans la théorie de l'apprentissage statistique. Ce principe leur confère un pouvoir prédictif élevé et leur permet d'éviter le problème de sur-apprentissage.

Machines à vecteurs support pour la régression (MVSR) permettent de convertir un problème de régression non linéaire en un problème linéaire en utilisant une projection des données de l'espace d'entrée vers un espace de description de dimension supérieure via une fonction noyau  $\phi$ . L'exemple de cette transformation est illustré dans la figure II.7 [46].



**Figure II.7.** Exemple de projection [47].

La relation non linéaire entre les entrées et les sorties dans le modèle des machines à vecteurs support peut être employée par la fonction de régression. Les résultats du modèle MVSR sont obtenus par l'équation suivante [48]:

$$f(x_i) = \omega^T \phi(x_i) + b, i = 1, 2, \dots, n \quad (\text{II.8})$$

$f(x_i)$  : La valeur prévue du modèle MVS.

$\phi(x_i)$ : La fonction non linéaire qui mappe l'espace fini-dimensionnel d'entrée dans un espace de dimension supérieure qui est implicitement créé.

$\omega$ : Le poids du modèle MVSR à optimiser.  $b$ : Le biais du modèle MVSR à optimiser.

La base de données a un vecteur d'entrée  $D$ -dimensionnel  $x_i \in R^D$  et une sortie scalaire  $y_i \in R$ .

Le modèle d'optimisation MVSR est donné par les équations suivantes (pour la base de données de formation) :

$$\left\{ \begin{array}{l} \min R(w, \xi, \xi^*, \varepsilon) = \frac{1}{2} \|w\|^2 + C \left[ v\varepsilon + \frac{1}{N} \sum_{i=1}^N (\xi_i + \xi_i^*) \right] \\ \text{subjective to : } y_i - w^T \phi(x_i) - b \leq \varepsilon + \xi_i \\ w^T \phi(x_i) + b - y_i \leq \varepsilon + \xi_i \\ \xi^*, \varepsilon \geq 0 \end{array} \right. \quad (\text{II.8})$$

$C$  : le paramètre utilisé pour équilibrer le risque empirique et le terme de complexité du modèle  $\|w\|^2$

$\xi_i^*$ : La variable de mou pour indiquer la distance de l'échantillon en dehors du tube  $\varepsilon$

En tant que problème standard d'optimisation non linéaire contraint, le problème ci-dessus, le problème ci-dessus peut être résolu en construisant le problème d'optimisation double basé sur les techniques de multiplicateurs Lagrange :

$$\left\{ \begin{array}{l} \max R(a_i, a_i^*) = \sum_{i=1}^N y_i (a_i, a_i^*) - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (a_i, a_i^*) (a_j, a_j^*) K(x_i, x_j) \\ \text{subjective to : } \sum_{i=1}^N y_i (a_i, a_i^*) = 0 \\ 0 \leq a_i, a_i^* \leq C/N \\ \sum_{i=1}^N (a_i + a_i^*) \leq C.v \end{array} \right. \quad (\text{II.10})$$

$K(x_i, x_j)$ : La fonction noyau qui satisfait l'état de Mercer.

$a_i$  and  $a_i^*$ : les multiplicateurs de Lagrange non négatifs, respectivement.

$$\hat{y} = f(x_i) = \sum_{i=1}^N (a_i - a_i^*) K(x - x_i) + b, i = 1, 2, \dots, n \quad (\text{II.11})$$

### **II.3.3. Forêt aléatoire (FA)**

Les forêts aléatoires, également connues sous le nom de forêt de décision aléatoire, sont des algorithmes d'apprentissage supervisé qui utilisent des méthodes d'apprentissage en ensemble pour résoudre des problèmes de régression et de classification. Elles fonctionnent en construisant un ensemble d'arbres de décision lors de l'apprentissage et en produisant la classe qui a la fréquence d'occurrence la plus élevée parmi les classes, ou en prédisant la moyenne (régression) des arbres individuels. Les forêts de décision aléatoire sont principalement conçues pour gérer le problème de sur ajustement associé aux arbres de décision [49].

Tim Kam Ho [50] a créé l'algorithme pionnier des forêts aléatoires en utilisant la méthode du sous-espace aléatoire. La méthode du sous-espace aléatoire est une manière d'appliquer l'approche de discrimination stochastique à la classification. Les forêts aléatoires utilisent la technique de bagging tout en exécutant les arbres en parallèle. Il n'y a pas d'interaction entre ces arbres lors de la construction, les forêts aléatoires veillent à ce qu'il n'y ait pas d'interaction entre les arbres. Avec quelques modifications bénéfiques, elles combinent le résultat de plusieurs prédictions en agrégeant de nombreux arbres de décision. Chacun de ces arbres extrait un échantillon aléatoire de l'ensemble de données principal lors de la génération de ses divisions. Cette action ajoute également un élément de hasard qui élimine le sur ajustement.

Les forêts aléatoires (FA) permettent de diviser le nombre de caractéristiques à chaque nœud en le limitant à un certain pourcentage du total avant l'apprentissage. Cette propriété hyper paramétrique garantit que le modèle FA ne dépend pas trop de chaque caractéristique individuelle.

#### **Avantages des forêts aléatoires**

- En comparaison avec d'autres algorithmes d'apprentissage, sa précision est considérée comme très élevée.
- Il est efficace pour traiter de grandes bases de données.
- Il accepte de multiples variables d'entrée sans n'en supprimer aucune.
- Il fournit des estimations des variables essentielles dans la classification ou la régression.
- À mesure que la construction de la forêt progresse, elle produit une approximation interne impartiale de l'erreur de généralisation.



- Il peut estimer efficacement les données manquantes tout en maintenant la précision lorsque la proportion de données manquantes est élevée.

### **Inconvénients des forêts aléatoires**

- On a constaté qu'elles avaient tendance à sur ajuster certains ensembles de données.
- Pour les données présentant un nombre différent de niveaux, elles sont observées comme biaisées en faveur des attributs ayant plus de niveaux.

### **II.3.4. Gradient Boosting Extrême (XGBoost)**

L'algorithme gradient boosting extrême (Extreme Gradient Boosting, XGBoost) est une variante du gradient boosting proposé par Chen et Guestrin [51], il est basé sur le principe du boosting. À la différence du bagging (le mot bagging est la contraction des mots Bootstrap et Aggregating), le boosting consiste à construire une famille de modèles (par exemple, des arbres CART avec une faible profondeur) de façon séquentielle où les apprenants faibles sont introduits de manière itérative. Chacun des modèles ajoutés à la combinaison, apparaît comme un pas vers une meilleure solution. Ce pas est franchi dans la direction opposée du gradient de la fonction de perte, lui-même approché par un arbre de régression. Ces modèles sont ensuite agrégés par une moyenne pondérée des estimations [52].

La fonction d'approximation obtenue par le modèle XGBoost à l'itération  $t$  est donnée par :

$$f_i^t = \sum_{k=1}^t f_k(x_i) = f_i^{t-1} + f_t(x_i) \quad (\text{II.12})$$

Avec  $f_t(x_i)$  est l'arbre de régression à l'itération  $t$ ,  $f_i^t$  et  $f_i^{t-1}$  sont les prédictions aux itérations  $t$  et  $t - 1$ .

Le modèle XGBoost optimise une fonction objective régularisée de la forme :

$$Obj^{(t)} = \sum_{k=1}^n l(\hat{y}_i, y_i) + \sum_{k=1}^n \Omega(f_i) \quad (\text{II.13})$$

avec  $l$  est la fonction de perte,  $n$  est le nombre d'observations et  $\Omega$  est le terme de régularisation défini par :

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \| \omega \|^2 \quad (\text{II.14})$$

où  $T$  est le nombre de feuilles au dernier niveau de l'arbre,  $\omega$  le score prédit par la feuille  $t$  ( $1 \leq t \leq T$ ),  $\gamma$  et  $\lambda$  sont des paramètres de régularisation.

Les hyperparamètres du modèle XGBoost sont : (max\_depth) : la profondeur maximale des arbres, (subsample) : rapport de sous-échantillon des exemples d'entraînement, (eta) : le pas de rétrécissement empêchant le surapprentissage et (n\_estimators) : nombre d'arbres du modèle XGBoost.

### II.4. Algorithmes d'analyse de séries temporelles

Bien que de nombreux travaux aient été réalisés en utilisant des modèles empiriques pour prédire les séries temporelles de l'irradiation solaire, il existe quelques inconvénients fondamentaux à ce type de modèles. Les données météorologiques requises par les modèles empiriques ne sont parfois pas disponibles dans les pays en développement. Un autre inconvénient est que la valeur cible de la prédiction est généralement de plusieurs heures à l'avance, voire plus. Cependant, la prédiction à court terme et en temps réel de l'irradiation solaire est nécessaire pour les systèmes modernes d'énergie renouvelable. Le modèle empirique de l'irradiation solaire est également incapable de détecter les changements anormaux ou soudains de l'irradiation solaire, ce qui est crucial pour le contrôle des systèmes basés sur l'irradiation solaire. Pour surmonter ces inconvénients, de nombreux chercheurs considèrent l'irradiation solaire comme une série temporelle et utilisent des approches d'analyse de séries temporelles pour le traiter. Nous allons maintenant examiner certaines de ces approches.

#### II.4.1. Autorégressif non linéaire avec entrées exogènes (ARNX)

Modèle autorégressif non linéaire avec entrées exogènes (Nonlinear autoregressive modèles with exogenous inputs ARNX) est couramment utilisé dans la modélisation de séries chronologiques. Il a une nature dynamique récurrente. ARNX comprend des connexions de rétroaction englobant plusieurs couches du réseau [53].

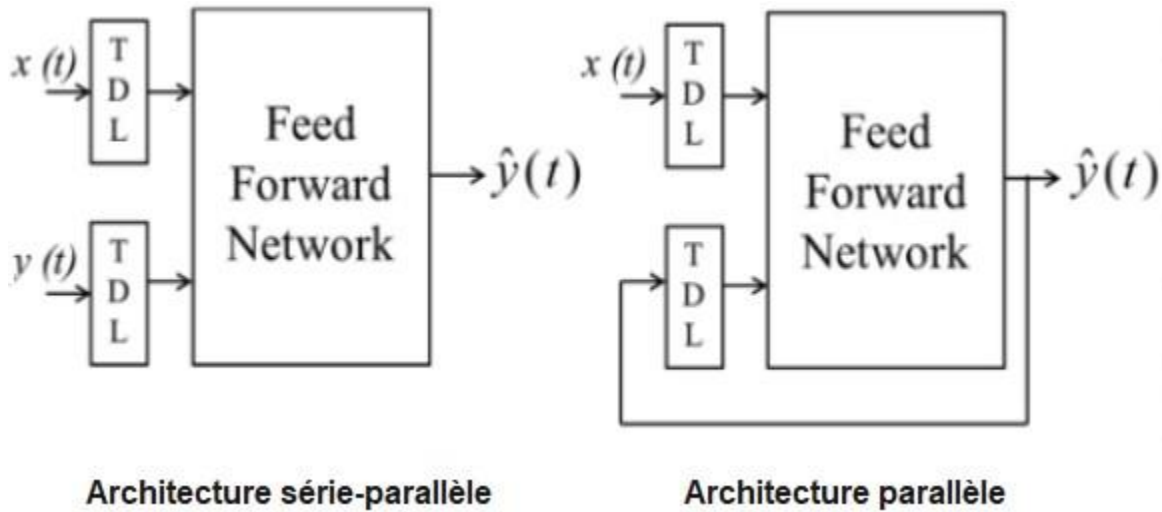
ARNX est composé de nœuds interconnectés, chaque nœud représente un neurone artificiel qui reçoit une ou plusieurs entrées et les élabore pour produire une sortie. Ces sommes passent par une fonction d'activation non linéaire. Les principaux avantages des réseaux ARNX par rapport aux autres approches RNA sont la convergence plus rapide pour atteindre les poids optimaux des connexions entre les neurones et les paramètres d'entrée et le nombre réduit de ces derniers pour calibrer et rendre le modèle efficace [54].

Afin d'obtenir les performances complètes du réseau de neurones ARNX pour la prédiction de séries temporelles non linéaires, il est intéressant d'utiliser sa capacité de mémoire en utilisant les valeurs passées des séries temporelles prédites ou vraies. Comme on peut le voir dans la figure II.8, il existe deux architectures différentes du modèle de réseau de neurones ARNX, l'architecture série-parallèle (également appelée boucle ouverte) et l'architecture parallèle (également nommée boucle fermée) donnée par les équations (II.6) et (II.7), respectivement :

$$\hat{y}(t + 1) = F(y(t), y(t - 1), \dots, y(t - n_y), x(t + 1), x(t), x(t - 1), \dots, x(t - n_x)) \dots \dots \dots \quad (II.15)$$

$$\hat{y}(t + 1) = F(\hat{y}(t), \hat{y}(t - 1), \dots, \hat{y}(t - n_y), x(t + 1), x(t), x(t - 1), \dots, x(t - n_x)) \dots \dots \dots \quad (II.16)$$

Où,  $F(\cdot)$  est la fonction de mappage du réseau de neurones,  $(t + 1)$  est la sortie du NARX au moment  $t$  pour le temps  $t + 1$ .  $\hat{y}(t), \hat{y}(t-1), \dots, \hat{y}(t - n_y)$  sont les sorties passées du NARX.  $y(t), y(t-1), \dots, y(t-n_y)$  sont les vraies valeurs passées de la série chronologique, également appelées valeurs de sortie souhaitées.  $x(t + 1), x(t), \dots, x(t - n_x)$  sont les entrées du NARX.  $n_x$  est le nombre de retards d'entrée et  $n_y$  est le nombre de retards de sortie.



**Figure II.8.** Différents architecture du réseau NARX [55].

**II.4.2. Autorégressif intégré moyenne mobile (ARIMA)**

Ce n'est qu'en 1927 que la notion de processus stochastique est proposée par Yule [56]. Elle voyait toute série chronologique naturelle comme la réalisation d'un processus stochastique. Sur la base de cette

idée, des modèles autorégressifs (AR) et moyens mobiles (MA) ont été développés. Depuis lors, de nombreux processus stochastiques ont été inventés et appliqués à l'analyse de séries chronologiques. Box et Jenkins ont intégré les processus stochastiques existants et développé le modèle ARIMA. En outre, ils ont également fourni une méthode étape par étape pour construire un modèle ARIMA approprié pour des séries temporelles spécifiques. Cela a fait d'ARIMA le modèle de prédiction de séries chronologiques le plus populaire [57].

Le modèle ARIMA est l'un des modèles univariés les plus populaires pour les séries temporelles. Sa popularité repose sur la célèbre méthodologie de Box-Jenkins, qui démontre que l'ARIMA est compétent pour modéliser différents types de séries temporelles. Il a été prouvé depuis longtemps que l'ARIMA est capable de représenter des séries temporelles présentant différentes propriétés statistiques. Cependant, il n'existait pas de méthode pour déterminer l'ordre approprié du modèle ARIMA jusqu'à ce que Box et Jenkins utilisent le coefficient d'autocorrélation et le coefficient d'autocorrélation partielle pour étudier l'ordre. Ils ont découvert qu'il existe une certaine relation statistique entre ces coefficients et l'ordre des processus autorégressifs (AR) et moyenne mobile (MA). Certains critères mathématiques sont également utilisés pour identifier l'ARIMA approprié. Le critère d'information bayésien (CIB) et critère d'information d'Akaike (CIA) et le sont les deux critères les plus utilisés. Ils sont tous les deux basés sur l'idée de minimiser l'erreur de prédiction à un pas en avant. L'erreur de prédiction finale d'Akaike (EPF) a été développée sur la base de l'CIA pour la prédiction à plusieurs pas en avant. Elle prend en compte l'erreur de prédiction sur plusieurs pas en avant et la pénalité pour le sur ajustement [58].

En plus de l'analyse de l'autocorrélation et des critères statistiques, certaines recherches montrent que des essais et erreurs sur des échantillons peuvent parfois donner une meilleure précision d'identification. Dans ce processus, l'ensemble de données d'apprentissage est divisé en plusieurs groupes, et une validation croisée est effectuée sur le modèle ARIMA avec différents ordres. Celui qui donne les meilleures performances est considéré comme le plus approprié [58].

Un autre problème dans le processus de construction du modèle concerne l'estimation des paramètres. L'estimation par maximum de vraisemblance est l'algorithme d'estimation le plus populaire en raison de sa simplicité et de son efficacité. En considérant les paramètres du modèle comme des variables aléatoires, Zellner a appliqué l'analyse bayésienne pour obtenir leur distribution. Certaines découvertes récentes soutiennent que la méthode d'estimation des paramètres corrigée de biais est meilleure que la méthode des moindres carrés [59].

Une méthode novatrice qui utilise un filtre à mémoire courte plutôt qu'un filtre à mémoire longue pour traiter les séries temporelles est présentée sous le nom de méthodologie ARARMA par Parzen [60]. Elle nécessite une étape d'identification des paramètres de plus que l'ARIMA traditionnel. Elle offre de meilleures performances que l'ARIMA, notamment en ce qui concerne les prédictions à long terme. Une autre stratégie pour réaliser des prédictions à plusieurs pas consiste à utiliser différents modèles AR pour effectuer la prédiction de séries temporelles avec des horizons différents. Cependant, cela augmenterait la difficulté de l'identification du modèle et de l'estimation des paramètres.

Une autre évolution basée sur l'ARIMA, visant à améliorer ses performances de prédiction à un pas, est l'ARIMA uni varié automatique. Il utilise un système expert pour générer et sélectionner le modèle ARIMA approprié, puis l'utilise pour effectuer la prédiction.

Lorsque l'entrée du modèle ARIMA contient plus d'une dimension, il est nécessaire d'adopter un modèle ARIMA sous forme de vecteur. Le modèle ARIMA avec une entrée vectorielle est appelé modèle ARIMA multi varié ou modèle VARIMA. Bien que cette idée ait été proposée pour la première fois dans les années 1950, il y a eu peu de recherche à ce sujet jusqu'aux années 1980 et 1990, lorsque les ordinateurs ont été utilisés dans le développement des modèles statistiques. Le modèle VARIMA repose sur les hypothèses d'exogénéité et de relation contemporaine, ce qui a posé de nouveaux défis aux chercheurs [61].

Certains chercheurs ont montré que les filtres de lissage pouvaient améliorer les performances du modèle VARIMA. En effet, les filtres de lissage peuvent réduire les fluctuations irrégulières présentes dans la série. Le modèle VARIMA peut théoriquement être utilisé pour effectuer des prédictions à la fois à long terme et à court terme. Pour explorer la relation entre la précision des prédictions et l'horizon de prédiction, certains chercheurs ont développé de nouvelles propriétés statistiques basées sur les prédictions cumulées à plusieurs pas en avant et les erreurs cumulées à plusieurs pas en avant. Cependant, très peu de travaux de recherche se sont penchés sur la manière de déterminer l'ordre du modèle VARIMA.

Un modèle d'auto régression qui prend un vecteur en tant que paramètre d'entrée est le modèle Vecteur Autorégression (VAR). On peut également considérer le VAR comme un VARIMA sans processus à moyenne mobile. Le VAR est principalement utilisé dans l'analyse des séries économiques, et il existe de nombreuses formes différentes de ce modèle. Cependant, l'un des inconvénients du VAR

est sa sensibilité au sur ajustement. Bien que son ajustement dans l'échantillon soit très bon, sa performance de prédiction hors échantillon peut parfois être assez médiocre. Pour surmonter cette lacune, Litterman et d'autres chercheurs ont imposé une distribution a priori sur les paramètres. Ils croyaient que les paramètres du modèle VAR étaient conformes à un certain comportement. Par exemple, ils étaient convaincus que les paramètres du modèle économique ressemblaient à une marche aléatoire. Cet algorithme est appelé modèle Bayesian Vecteur Auto régression (BVAR) [62].

Sur la base des algorithmes précédents, Engle et Granger ont développé un nouveau concept appelé les modèles de correction d'erreur- Error Correction Models (ECM). De nombreuses expériences ont montré que les performances des ECM sont meilleures que celles des modèles VAR et BVAR [63]. Surtout en ce qui concerne les prédictions à long terme, les ECM peuvent surpasser les modèles VAR et BVAR. Après avoir introduit l'approche bayésienne dans l'estimation des paramètres des ECM, ces modèles peuvent présenter des résultats encore plus précis qu'auparavant.

Hokoi et al. [64] ont utilisé un modèle ARMA pour développer un modèle stochastique des séries temporelles de rayonnement solaire horaire pour les mois d'été. Après la phase de transformation, il a été constaté qu'un modèle ARMA (3, 3) présentait les meilleures performances. La fonction d'autocorrélation des données réelles et des données simulées coïncidait à de petits retards temporels. Cependant, à de grands retards temporels, le modèle proposé ne parvenait pas à suivre les données réelles.

Chowdhury et Rahman [65] ont proposé un algorithme pour utiliser des données horaires dans la prévision du rayonnement solaire. Ils ont séparé l'ensemble de données en rayonnement solaire en ciel clair et en transmissives nuageuses. Pour le ciel clair, les auteurs ont supposé que le modèle empirique existant était suffisant pour effectuer la prévision. Ils ont utilisé un modèle ARMA pour la prévision du rayonnement solaire par temps nuageux. Ils ont déclaré que ces modèles étaient assez précis, sauf en cas de changements de transition dans la couverture nuageuse devant le soleil. Un travail similaire a été proposé par Mora-Lopez et Sidrach-de-Cardona [66]. Leur travail utilise un modèle ARIMA (0, 1, 1) pour la prédiction de l'irradiation solaire global en pourcentage de la valeur maximale disponible.

### II.5. Revue des travaux sur la modélisation de l'irradiation solaire

Plusieurs modèles ont été proposés dans la littérature pour la prédiction de l'irradiation solaire, tels que les modèles basés sur l'intelligence artificielle (IA), les modèles empiriques, les séries temporelles et les méthodes hybrides, les modèles développés à partir des images satellitaires, cela dans différents

endroits du monde. Parmi ces méthodes, les deux modèles les plus couramment utilisés sont les réseaux de neurones artificiels de type Perceptron à Multicouches (RNA-PMC) et la Machine à vecteur support (MVS). En revanche, le modèle de la Forêt aléatoire (FA) et les algorithmes d'analyse de séries temporelles sont les moins utilisés.

Jiang [67] a développé des modèles de réseau neuronal comportant une couche cachée avec 7 neurones pour huit villes en Chine, et les a comparés à six modèles empiriques pour une prédiction mensuelle moyenne quotidienne. L'algorithme de propagation arrière à alimentation directe a été appliqué dans cette analyse. Les données de rayonnement solaire global enregistrées pendant dix ans, de 1995 à 2004, ont été utilisées. Les variables d'entrée utilisées dans tous les modèles sont la durée d'ensoleillement, la latitude  $\lambda$  ( $^{\circ}$ ) et l'altitude  $h$  (km). Les critères d'évaluation utilisés sont le pourcentage d'erreur moyenne (PEM), l'erreur absolue moyenne (EAM) et l'erreur quadratique moyenne (EQM), et ont montré que les modèles de Réseau Neuronal Artificiel (ANN) donnent de meilleurs résultats que les modèles empiriques.

Fortin et al.[68] ont comparé trois modèles classiques avec les mêmes modèles optimisés en utilisant un schéma non linéaire non contraint standard et les modèles de réseaux de neurones à couches multiples. Trois années de données (2004-2006) provenant de 11 sites au Canada sont utilisées pour les tests et la validation des trois catégories de modèles. Les meilleures performances ont été obtenues avec un PMC de 8 neurones cachés.

Dahmani et al. [69] ont testé avec succès la capacité de PMC à générer des estimations d'irradiance oblique à très court terme (5 min) à Alger (Algérie). Données d'entrée utilisées : rayonnement global horizontal, rayonnement extraterrestre, déclinaison, zénith et azimut. Des études ont montré que l'élimination de l'azimut comme paramètre d'entrée peut améliorer la précision du modèle.

Khatib et al. [70] ont utilisé quatre types de modèles de réseaux neuronaux pour prédire à la fois le rayonnement solaire horaire et le rayonnement solaire diffusé. Réseau neuronal Feed Forward (RNFF), cascade réseau de neurones avancés, réseau de neurones Elman ELMNN), et GRNN devaient déterminer la production de rayonnement solaire à Kuala Lumpur City. Plusieurs variables métrologiques ont été utilisées pour obtenir les résultats visés (longitude, latitude, température, rapport d'ensoleillement, humidité, mois, jour et heure). Les résultats ont montré que le prédicteur GRNN était meilleur que d'autres modèles, où la topologie ELMNN montrait les pires performances.

Chen et al. [71] ont combiné les deux modèles de réseaux neuronaux avec un modèle logique flou pour prédire le rayonnement solaire global horaire et quotidien. Trois variables ont été choisies comme attributs d'entrée, y compris les conditions du ciel (ensoleillé, nuageux et pluvieux), le temps et la température. Les résultats indiquent que le modèle neuronal flou proposé a atteint une bonne précision dans différentes conditions météorologiques.

Kaushika et al. [72] ont fait une étude comparative des estimations du modèle réseaux neurone artificiel (RNA) avec la météorologie de base de données. Les résultats ont montré une bonne compatibilité. Le rayonnement solaire global sur les plans inclinés a été étudié en utilisant des conditions atmosphériques isotropes et anisotropiques. Les résultats semblent favoriser un modèle isotrope tout au long de l'année.

Laidi et al. [73] ont utilisé des données appartenant à 13 sites dans différentes zones climatiques en Algérie pour estimer l'irradiation inclinée. Huit paramètres d'entrée ont été appliqués : la latitude du lieu, la température moyenne, l'humidité relative, le facteur de turbidité de Linke, le coefficient d'Angstrom, le rayonnement extraterrestre, le rayonnement global horizontal et l'angle zénithal. La configuration optimale comprenait 35 neurones dans la couche cachée.

Mehdizadeh et al. [74] ont utilisé trois modèles différents, à savoir Gene Expression Programming (GEP), RNA et ANFIS, pour le rayonnement solaire global quotidien à Karmen en Iran. Les meilleurs résultats de l'étude pertinente ont été observés dans le modèle ANN. Dans ce modèle,  $R^2$  a été calculé à 0,935.

Marzo et al. [75] ont tenté de prédire le rayonnement solaire global quotidien de 13 différentes stations. Les auteurs n'ont utilisé RNA que comme algorithme d'apprentissage automatique dans cette étude. Dans l'étude, le rayonnement solaire extraterrestre, la température minimale et la température maximale ont été utilisés pour former l'algorithme RNA. Dans les résultats,  $nREQM = 13 \%$ ,  $nEBM = 4 \%$  et  $R = 0,800$  ont été calculés comme les meilleurs résultats.

Quej et al. [76] ont prédit des données quotidiennes sur l'irradiation solaire globale de six stations au Mexique en utilisant trois algorithmes d'apprentissage automatique, à savoir MVS, RNA et ANFIS. Les auteurs ont utilisé l'irradiation solaire extraterrestre, les précipitations, la température minimale et les données de température pour l'apprentissage des algorithmes. Dans l'étude pertinente, les meilleurs résultats ont été obtenus en MVS avec  $REQM = 2,578$ ,  $EAM = 1,97$  et  $R^2 = 0,689$ .



Mohammad et al. [77] ont utilisé trois structures différentes du réseau neuronal MLP-BP pour prédire le rayonnement solaire quotidien moyen. Quatre variables d'entrée ont été utilisées dans ces neurones d'entrée (température ambiante quotidienne maximale, durée d'ensoleillement, humidité et vitesse du vent). Les résultats ont révélé que l'utilisation de neuf neurones cachés dans des couches cachées fonctionnait mieux que d'autres structures.

Une étude comparative a été menée par Benali et al. [78] Prédiction du rayonnement horizontal global et de ses deux composantes directe et diffuse avec un horizon temporel allant d'une heure à 6 heures à l'avance sur le site d'Odello (France). Autrement dit, les modèles de persistance intelligents, les modèles RNA et les modèles FA sont utilisés comme prédicteurs. Il a été démontré que le modèle FA fonctionne mieux lorsque l'horizon de prédiction est augmenté par rapport au modèle RNA. Les résultats montrent que les prévisions Le printemps et l'automne ne sont pas aussi fiables que l'hiver et l'été.

Sharadga et al. [79] ont utilisé plusieurs méthodes de prédiction de séries chronologiques et des méthodes statistiques pour prédire la production d'énergie photovoltaïque des centrales photovoltaïques à grande échelle. Quatre modèles différents ont été adoptés pour mettre en œuvre les modèles de prédiction, soit le modèle de série chronologique de réseau neuronal artificiel (RNA), le modèle de moyenne mobile autorégressive (ARMA), l'ARIMA et le modèle de série chronologique de moyenne mobile intégrée autorégressive saisonnière (SARIMA). Huit délais de temps sont appliqués et utilisés comme variables d'entrée pour prédire la puissance de sortie PV. Les résultats ont montré que les réseaux neuronaux sont plus précis que d'autres modèles lorsqu'ils sont utilisés pour la prédiction de séries chronologiques de puissance PV et nécessitent moins de temps de calcul. En outre, les prédicteurs développés étaient efficaces pour prédire la puissance produite des centrales photovoltaïques pour seulement une heure à l'avance sans mesures supplémentaires de l'irradiance solaire ou des conditions météorologiques. Par conséquent, on constate que les prévisions des séries chronologiques pour les centrales photovoltaïques ne sont fiables qu'une heure à l'avance.

Kumar et al. [80] ont comparé les modèles de réseaux de neurones artificiels (Artificial Neural Network, ANN) aux modèles de régression linéaire pour estimer le rayonnement solaire global mensuel. Les résultats de cette recherche ont révélé que les modèles ANN ont surpassé les méthodes de régression.

Fan et al. [52] ont comparé les méthodes SVR, XGBoost et quatre modèles empiriques. Les deux modèles SVR et XGBoost ont montré une capacité prédictive comparable et ont surpassé les modèles empiriques. De plus, le modèle XGBoost était le plus efficace en termes de stabilité et de coût de calcul.

Ismail et al. [81] ont construit un modèle de moyenne mobile intégrée autorégressive (ARIMA) pour prévoir le rayonnement solaire quotidien intra-échantillon et hors-échantillon dans onze districts de la péninsule malaisienne. Les données quotidiennes sur l'irradiation solaire utilisées dans cette recherche sont recueillies par le National Aeronautique and Space Administration (NASA) pour les années 2009 à 2019. L'idée principale de cette étude est de construire un modèle intelligent qui peut prédire les données des séries chronologiques actuelles utilisées dans la recherche (dans l'échantillon) et les données en dehors de la plage des séries chronologiques actuelles (hors échantillon). Tous les incidents d'insolation sur une surface horizontale sont considérés pour recueillir la quantité quotidienne moyenne du rayonnement solaire total. Les résultats ont prouvé que le modèle ARIMA est capable de modéliser et de prévoir les données sur le rayonnement solaire.

Bounoua et al. [82] ont estimé le rayonnement solaire global quotidien à l'aide de plusieurs méthodes empiriques et d'apprentissage automatique. Vingt-deux modèles empiriques, des techniques de réseaux neuronaux artificiels (RNA) et des méthodes d'ensemble basées sur des arbres ont été testé et examiné pour estimer le rayonnement solaire global quotidien dans cinq sites marocains. Les mesures solaires et météorologiques quotidiennes des cinq sites marocains ont été recueillies et utilisées pour développer les modèles de prédiction sélectionnés. Plusieurs combinaisons de variables ont été utilisées pour mettre en œuvre les modèles empiriques. Alors que la température moyenne quotidienne de l'air, l'humidité relative moyenne quotidienne, la direction et la vitesse du vent, la pression quotidienne de la vapeur d'air déficient et le gradient de température quotidien sont les variables d'entrée adoptées pour prédire le rayonnement solaire global quotidien à partir de modèles intelligents. Les résultats ont révélé que le modèle empirique des facteurs de température et de géographie qui exige les mesures quotidiennes de la température de gradient et les paramètres du jour de l'année ont obtenu la précision optimale par rapport à d'autres cas tous considérés. Pour les modèles d'apprentissage automatique, on constate que tous les arbres de décision, les réseaux neuronaux et la forêt aléatoire ont donné de bonnes performances, où la forêt aléatoire a mieux fonctionné et plus robuste que d'autres modèles intelligents.

Sridharan [83] a développé trois modèles : GRNN (réseau neuronal de régression généralisée), ANN (réseau neuronal artificiel) et Fuzzy (flou) pour la prédiction mensuelle de l'irradiation solaire en utilisant

des paramètres météorologiques. Les données d'entrée prises pour ses modèles sont : le mois de l'année, l'altitude, la longitude, la latitude, la température ambiante et la durée moyenne d'ensoleillement pour quatre villes (Jodhpur, Shillong, New Delhi et Kolkata) en India. L'évaluation des modèles montre que l'erreur relative moyenne absolue (nEMA) pour l'utilisation de GRNN, de la logique floue et du réseau neuronal artificiel est de 3,55 %, 4,64 % et 5,49 % respectivement.

Mellit et al. [84] ont développé un modèle hybride permettant de générer des irradiances solaires globales horizontales en combinant un RNA et une librairie de matrices de transition de Markov à partir de la latitude, longitude et altitude du lieu. Cette étude a été réalisée sur la base de 60 stations en Algérie sur la période 1991-2000. La précision est bonne puisque l'Erreur Quadratique Moyenne (REQM) quel que soit le site n'excède pas 8% et permet d'obtenir rapidement et avec un nombre réduit de paramètres des données d'irradiances solaires journalières en des lieux ne disposant pas de mesures

Yacef et al. [85] ont développé un modèle empirique combiné et un réseau neuronal bayésien (RNB) pour estimer l'irradiation solaire global quotidien sur la surface horizontale de Ghardaïa, en Algérie. Les coefficients du modèle empirique ont été estimés à l'aide des bases de données de l'irradiation global, des températures maximales et minimales en 2006. 6 mois en 2007 (été : mai, juin, juillet, hiver : octobre, novembre, décembre) pour tester les modèles de calage ; 6 mois en 2012 (février-juillet) pour la validation Généralité du modèle développé et du modèle RNB. Les résultats ont montré que le nouveau modèle calibré était capable d'estimer très précisément l'irradiation solaire globale.

Nwokolo et al. [86] ont réalisé une évaluation numérique de la littérature mondiale sur l'énergie solaire. En utilisant les modèles suivants, basés sur la lumière du soleil, la température, les précipitations, le noyau nuageux, l'humidité comparative et les paramètres hybrides, la disponibilité des 356 modèles empiriques et des 68 formes fonctionnelles a été acquise.

Ayodele et Ogunjuyigbe [87] ont montré la performance d'évaluation de la modélisation empirique pour les modèles de modélisation du rayonnement solaire mondial à Ibadan (Nigeria). Ces modèles ont été dérivés de trois modèles de base : Angström–Prescott, Hargreaves–Samani et Gracia. Les paramètres d'entrée sont les suivants : rayonnement solaire moyen global, température maximale de l'air par jour, ensoleillement quotidien moyen et température minimale de l'air par jour pendant 9 ans (2000-2008).

Benmouiza et Cheknane [88] ont prédit l'irradiation solaire global horaire à l'aide d'une combinaison de modèle de séries chronologiques de réseaux neuronaux autorégressifs (NRA) à moyenne k et non

linéaires. Pour valider les résultats, Les tests ont utilisé trois grappes, 25 retards et 13 neurones. Les résultats recueillis sont comparés avec des moyennes autorégressives et mobiles (ARMA) naïves et d'autres modèles dans le field. Les résultats ont indiqué que l'utilisation du cluster k-means avec NAR améliorera la performance des prévisions.

Piri et al. [89] ont comparé les algorithmes SVR avec les méthodes empiriques traditionnelles pour prédire l'irradiation solaire dans deux sites en Iran. Les résultats de cette recherche ont montré que les modèles SVR étaient les plus adaptés pour estimer l'irradiation solaire quotidienne dans la région subtropicale humide de Chine.

Suyono et al. [90] ont prédit l'intensité du rayonnement solaire "Indonésie" à l'aide de la méthode Extreme Learning Machine (ELM) et de la Régression Linéaire Multiple (MLR). Ils ont utilisé la température, la durée d'ensoleillement et l'irradiation solaire comme données d'entraînement. Les données ont été collectées sur une période de cinq ans, et leurs résultats ont montré qu'ELM s'est nettement mieux comportée que la MLR conventionnelle, avec un RMSE de 15,29 % et un MAE de 24,79 %. Le tableau 4.1 montre les performances prédictives de quelques algorithmes d'intelligence artificielle utilisés dans la littérature pour estimer l'irradiation solaire.

## Chapitre II : Synthèse bibliographique sur la modélisation de l'irradiation solaire par les modèles d'intelligence artificielle

**Tableau II.2.** Exemples significatifs de l'application des différents types publiés à la prédiction de l'irradiation solaire.

Références	Sites	Modèle	Indicateur d'évaluation	Intervalles des données
Jiang el. [67]	China	RNA	$R^2 = 0.98$ REQM = 1.228	1995 à 2004
Fortin et al. [68]	Canada	PMC	EAM = 24.5% REQM = 21.3%	2004 à 2006
Benghanem and Mellit [91]	Arabie saoudite	RBF	R = 98.80% REQM = 0.013	1998 à 2002
Linares et al. [92]	Espagne	BP – FF- PMC	R = 94 REQM=14.29 à 16.4%	1999 à 2009
Koca et al. [93]	Turquie	BP – FF- PMC	$R^2 = 97.86$ à $99.87\%$ REQM = 3.58 à 6.9%	2005 à 2008
Voyant et al. [94]	France	RNA	REQM=23.4%	--
Piri et al. [89]	Iran	SVR et méthodes empiriques	R = 0,9330 RMSE=0,4515 kWh/m <sup>2</sup> /jour	7 ans
Siham et al.[95]	Alegria	RNA	REQM = 4.71% EBM = 0.021%	2013
Fan et al. [52]	China	SVR, méthodes empiriques	$R^2 = 0,7760$ REQM=1,002 kWh/m <sup>2</sup> /jour	1966 à 2015
Benali et al. [78]	France	RNA	REQM = 22.57%	--
Castangia et al. [96]	Italie	LSTM RF	$R^2 = 0.93$ $R^2 = 0.97$	2010 à 2015
Huang et al. [97]	États-Unis	WPK CNN-LSTM-PMC	nREQM = 23.917 nREQM = 17.4132 nREQM = 15.479	2009 à 2016
Ismail et al. [81]	Malaysia	ARIMA	REQM = 22.57%	2009 à 2019
Bounoua et al. [82]	Maroc	RNA RF	R= 87.53–96.20% nEAM = 5.84–11.81% nREQM = 7.85–15.33%	2011 à 2015
Guermoui et al. [98]	Alegria	MVS	RMSE = 11.35% R= 12.61%	2013 à 2015
Bamisile [99]	Nigeria	RNA	R = 0.9546, RMSE= 82.22 W/m <sup>2</sup>	2005 à 2016
Üstün et al. [100]	Turkey	SMGRT	$R^2 = 0.960$ MBE = 0.156 REQM = 1.371	2007 à 2009 et 2011 à 2016
Youness El Mghouc [101]	Maroc	RNA	R = 0.987 EMAP = 1.294	2005 à 2022
Sridharan [83]	India	RNA	nEAM = 3.32	---

## II.6. Performance et évaluation des prédicteurs

Afin de mesurer le pouvoir prédictif des différents modèles AI supervisé, plusieurs métriques ont été utilisées dans la littérature [102], [103] :

**Coefficient de détermination  $R^2$**  : Il est défini comme la proportion de variance expliquée par rapport à la variance totale, il permet de mesurer l'adéquation entre les valeurs mesurées  $Y_i$  et estimées  $\tilde{Y}_i$  par le modèle étudié. Il est donné par la formule :

$$R^2 = 1 - \frac{\sum_{i=1}^n (\tilde{Y}_i - \tilde{Y}_{im})^2}{\sum_{i=1}^n (Y_i - Y_{im})^2} \quad (\text{II.17})$$

avec  $Y_{im}$  la valeur moyenne des valeurs mesurées et  $\tilde{Y}_{im}$  la valeur moyenne des valeurs estimées.

**Coefficient de corrélation  $R$**  : Il représente la racine carrée du coefficient de détermination, il permet de quantifier la corrélation entre les valeurs prédites par le modèle et celles mesurées.

**L'erreur de biais moyenne (EBM)** : Il est défini comme étant l'écart algébrique moyen entre les valeurs prédites  $\tilde{Y}_i$  et les valeurs mesurées  $Y_i$ , sa formule est :

$$\mathbf{EBM} = \frac{1}{n} \sum_{i=1}^n (\tilde{Y}_i - Y_i) \quad (\text{II.18})$$

Une valeur positive de MBE signifie que le modèle surestime les mesures réelles alors qu'une valeur négative signifierait l'inverse.

**L'erreur absolue moyenne (EAM)** : Elle indique l'écart absolu moyen entre les valeurs prédites et celles mesurées. Elle a l'avantage de faire disparaître la compensation des erreurs, sa formule est donnée par :

$$\mathbf{EAM} = \frac{1}{n} \sum_{i=1}^n |\tilde{Y}_i - Y_i| \quad (\text{II.19})$$

**Erreur moyenne absolue en pourcentage (EMAP)** : L'EMAP représente la précision sous forme de pourcentage d'erreur. Étant exprimé en pourcentage, ce chiffre pourrait être plus intuitif à saisir que les autres mesures statistiques. Elle est donnée par la relation:

$$\mathbf{EMAP} = \frac{100}{n} \left( \sum_{i=1}^n \left| \frac{\tilde{Y}_i - Y_i}{Y_i} \right| \right) \quad (\text{II.20})$$

**Racine de l'erreur quadratique moyenne (REQM):** Elle représente la racine de la moyenne arithmétique des carrés des écarts entre les prédictions  $\tilde{Y}_i$  et les expérimentales  $Y_i$ . Elle est donnée par la relation :

$$\text{REQM} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\tilde{Y}_i - Y_i)^2} \quad (\text{II.21})$$

**Erreur standard de prédiction (ESP):** Est une mesure qui évalue la dispersion des valeurs prédites par rapport aux valeurs réelles dans un modèle de régression. Elle représente la précision des prédictions du modèle par rapport aux données observées.

$$\text{ESP} = \frac{\text{REQM}}{Y_{im}} \times 100 \quad (\text{II.22})$$

## II.7. Conclusion

Dans ce chapitre, nous avons présenté une brève introduction à l'intelligence artificielle et aux algorithmes des séries temporelles pour la prédiction de l'irradiation solaire. Cette évaluation s'avérerait précieuse pour guider le choix du modèle le plus approprié et précis, en se basant sur les données météorologiques disponibles. Dans le cadre de cette étude, nous avons entrepris une analyse statistique approfondie et effectué une classification des modèles de l'irradiation solaire existants. La revue exhaustive de la littérature que nous avons menée nous a permis de dégager les principales étapes de l'élaboration d'un modèle d'estimation ou de prédiction de l'irradiation solaire :

- Collecte de données provenant de multiples localités.
- Le choix de la technique de modélisation repose sur la qualité des données à disposition.
- Exploration des corrélations entre divers paramètres météorologiques et l'irradiation solaire.
- Sélection de la structure des modèles.
- Choix des paramètres prédominants en fonction de leur disponibilité et de leur corrélation avec l'irradiation solaire.
- Test du modèle en termes d'indépendance et de cohérence.

# Chapitre III

Évaluation de l'efficacité de l'apprentissage amélioré dans la  
prédiction de l'irradiation horizontale globale



### **III.1. Introduction**

Les préoccupations croissantes concernant les effets du changement climatique et la nécessité de diversifier les sources d'énergie ont entraîné une augmentation significative du développement des sources d'énergie renouvelables. Parmi celles-ci, l'énergie solaire est devenue une option viable en raison de son abondance et de son potentiel de réduction des émissions de carbone. De plus, comme les ressources pétrolières et gazières sont de moins en moins disponibles, le développement de sources d'énergie renouvelable devient de plus en plus important pour la sécurité énergétique à long terme du pays. Selon le récent rapport du groupe intergouvernemental sur le changement climatique (GIEC), l'énergie solaire a le potentiel de répondre à une partie importante des besoins énergétiques mondiaux, et l'Algérie ne fait pas exception [104].

L'énergie solaire est au centre de la politique énergétique ambitieuse de l'Algérie, qui alloue des ressources importantes aux énergies solaires thermiques et photovoltaïques. Les projections indiquent que l'énergie solaire représentera plus de 37 % de la production d'électricité du pays d'ici 2030 [105]. La durée annuelle d'ensoleillement en Algérie est de plus de 3900 heures au Sahara et 3000 heures sur les plateaux. Le gain d'énergie quotidien sur une surface horizontale de 1 m<sup>2</sup> est en moyenne de 5 KWh. Dans la plupart des régions du pays [103]. Ainsi, l'énergie solaire est une bonne base pour aider le pays à répondre à ses besoins énergétiques. Cependant, des prédictions précises de l'irradiation solaire sont nécessaires pour exploiter pleinement ce potentiel.

Pour relever ces défis, les chercheurs ont mis au point des modèles qui utilisent des données météorologiques facilement disponibles pour prédire plus précisément l'irradiation solaire globale. Ces modèles prédictifs sont de plus en plus avancés chaque jour, mais les résultats varient selon l'emplacement. Par conséquent, il est important d'utiliser des techniques sophistiquées de prédiction l'irradiation solaire globale pour améliorer la précision de prédiction du potentiel d'énergie solaire en Algérie[106].

L'objectif principal de cette contribution est de développer une méthode pour optimiser les paramètres de l'apprentissage automatique traditionnel en utilisant les réseaux de neurones artificiels type de Perceptron à Multicouches (RNA-PMC) et Machine à vecteur support (MVS) augmentant ainsi la fiabilité des prédictions horaires de l'irradiation horizontale globale dans les régions de Bouzaréah et

Tamanrasset. Les résultats de la prédiction des irradiations horizontale globale ont fait objet d'une publication dans le journal " Computers, Materials & Continua " (Dahmani et al. [107]). Dans ce chapitre, les résultats obtenus sont discutés.

### **III.2. Région étudiée et collecte de données**

Dans ce travail, deux stations radiométriques ont été utilisées pour compiler la base de données. La première station, "Shems", située à Bouzaréah en Algérie, a enregistré des données expérimentales en utilisant des pyranomètres « Kipp et Zonen » pour mesurer l'irradiation horizontale globale (IHG). La deuxième station, située à Tamanrasset dans le désert du Sahara au sud de l'Algérie, est équipée d'un pyranomètre « Eppley PSP » et dispose des plus grandes ressources en énergie solaire dans un environnement désertique aride. Tableau III.1 présente des informations complètes sur les deux stations utilisées à des fins de l'apprentissage et de test, y compris le nom de la station, la latitude, la longitude, l'altitude, la zone climatique et les périodes de données. Le tableau indique que la station de Bouzaréah, est située dans la zone climatique méditerranéenne et couvre la période entre le 1<sup>er</sup> janvier à 31 décembre 2014. En revanche, la station de Tamanrasset, se trouve dans la zone climatique du désert chaud et couvre la période du 1<sup>er</sup> juillet 2019 au 17 décembre 2020. Les informations dans le tableau sont cruciales pour comprendre les données utilisées dans l'enquête et interpréter les résultats obtenus.

**Tableau III.1.** Région géographique et période couverte par les deux stations.

<b>Station ID</b>	<b>Latitude [°]</b>	<b>Longitude [°]</b>	<b>Altitude [m]</b>	<b>Zone climatique</b>	<b>Périodes des données</b>
<b>Bouzaréah</b>	36.80	3.17	357	Climat méditerranéen	01.01.2014 à 31.12.2014
<b>Tamanrasset</b>	22.78	5.51	1378	désert chaud	01.07.2019 à 17.12.2020

Le tableau III.2 présente les résultats de l'analyse statistique des variables d'entrée et de sortie de la prédiction de l'irradiation horizontale globale aux stations Bouzaréah et Tamanrasset pour la température, l'humidité, la vitesse de vent et l'irradiation horizontale globale. La station de Bouzaréah enregistre une température maximale de 44,11 °C, une humidité minimale de 8,28 %, une vitesse moyenne du vent de 4,62 m/s et un IHG moyen de 517,88 Wh/m<sup>2</sup>. La station de Tamanrasset enregistre une température maximale de 38,5 °C, une humidité minimale de 2 %, une vitesse moyenne du vent de 5,19 m/s et une l'irradiation horizontale globale moyenne de 678,76 Wh/m<sup>2</sup>. La corrélation entre ces

variables et l'irradiation horizontale globale est examinée plus en détail dans le tableau III.3. Les résultats montrent que la température est positivement corrélée à l'irradiation horizontale globale aux deux stations, l'humidité ayant une corrélation négative à Bouzaréah et une faible corrélation négative à Tamanrasset. La vitesse du vent a une faible corrélation négative à Bouzaréah et une corrélation positive à Tamanrasset. Ces corrélations peuvent être utiles pour développer des modèles précis de prédiction de l'irradiation horizontale globale et concevoir des systèmes efficaces d'énergie solaire.

**Tableau III.2.** Analyse statistique des variables d'entrée et de sortie pour la prédiction de l'irradiation horizontale globale à deux stations.

Station	Statistique	Température [°C]	Humidité [%]	Vitesse du vent [m/s]	Irradiation horizontale globale [Wh/m <sup>2</sup> ]
<b>Bouzaréah</b>	Max	44.11	92.21	14.24	1027.00
	Min	5.56	8.28	0.10	120.25
	Moyenne	24.67	40.55	4.63	517.89
	Std	7.79	16.61	2.39	244.72
<b>Tamanrasset</b>	Max	38.50	99.00	13.90	1293.54
	Min	1.50	2.00	0.00	35.50
	Moyenne	26.86	19.70	5.19	678.76
	Std	7.65	10.06	2.71	300.83

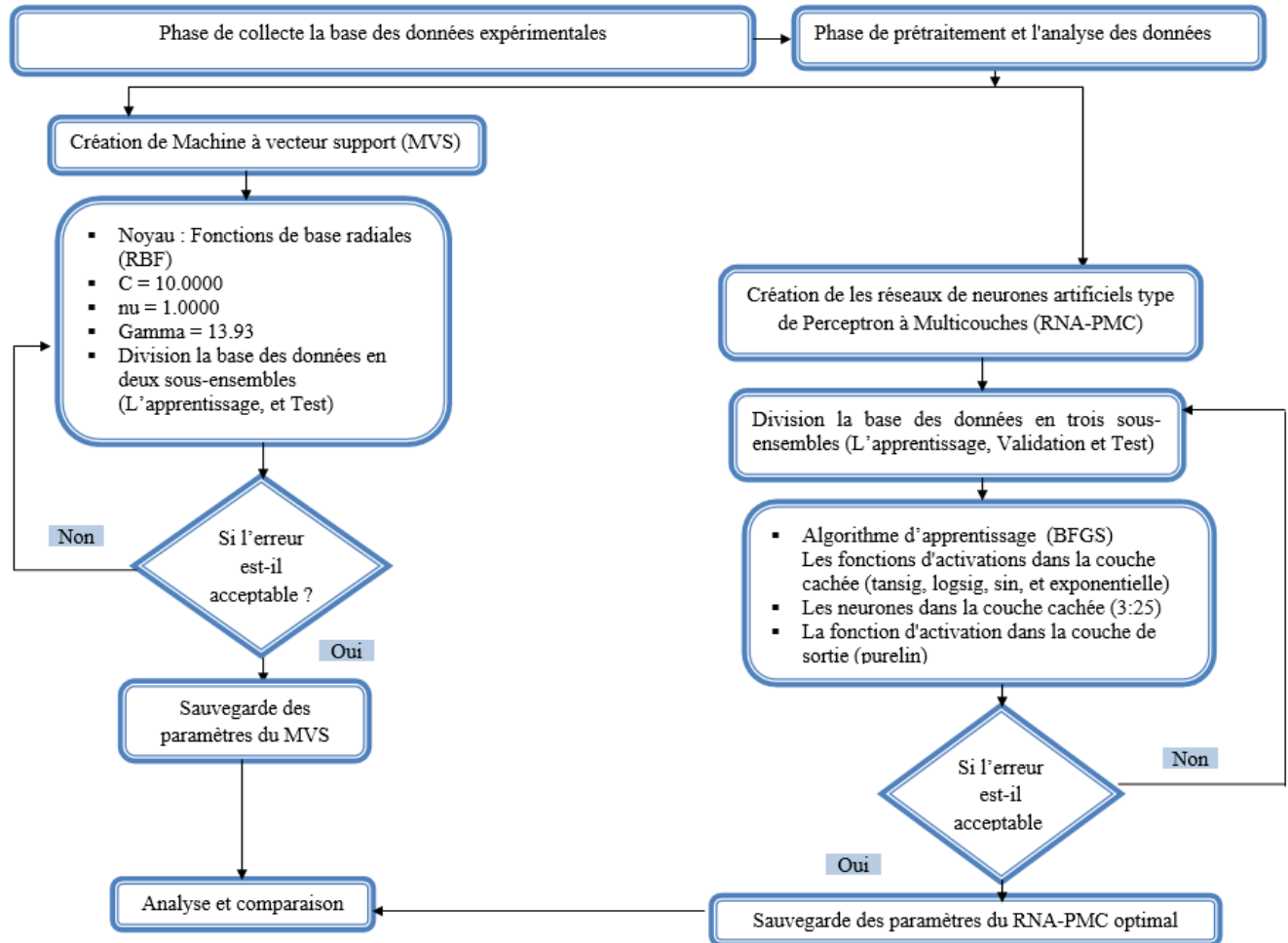
**Tableau III.3.** Corrélations climat-sortie pour la prédiction de l'irradiation horizontale globale à deux stations.

Station	Température [°C]	Humidité [%]	Vitesse du vent [m/s]	Direction de Vent[°]	Pression [mbar]
<b>Bouzaréah</b>	0.512	-0.559	-0.093	0.129	-0.1567
<b>Tamanrasset</b>	0.448	-0.298	0.163	-	-

### III.3. Evaluation et développement du modèle

Deux modèles ont été utilisés pour élaborer une prédiction précise de l'irradiation horizontale globale horaire : le (RNA-PMC) et le MVS. Le processus utilisé pour évaluer et améliorer la structure des modèles (RNA-PMC) et MVS est présenté en détail à la figure III.1. La base de données (BD) a été fractionnée en plusieurs ensembles distincts pour chaque modèle. Concernant le modèle RNA-PMC, la BD a été répartie aléatoirement en trois sous-ensembles : 80 % pour la phase d'apprentissage, 10 % pour la phase de validation, et 10 % pour la phase de test. Quant au modèle MVS, la même base de données a

été aléatoirement divisée en deux sous-ensembles : 80 % pour la phase d'apprentissage et 20 % pour la phase de test.



**Figure III.1.** Diagramme de la méthodologie pour calculer les deux modèles (RNA-PMC) et (MVS).

Plusieurs techniques ont été utilisées pour déterminer le modèle de RNA-PMC le plus efficace. L'algorithme d'apprentissage BFGS quasi-Newton (trainbfg) a été utilisé, quatre fonctions d'activation (log sigmoïde, tangentes hyperboliques, exponentielle et sin) ont été utilisées dans la couche cachée, et une seule fonction de transfert (identité) a été utilisée.

Le nombre de neurones dans la couche cachée a été varié de 3 à 25 neurones, afin d'obtenir le modèle le plus précis possible. Un processus de mise à l'essai itératif a été effectué pour déterminer le modèle de RNA-PMC offrant le meilleur modèle optimal. Un modèle MVS optimal a été développé en utilisant la stratégie d'apprentissage automatique machine à vecteur de support (MVS). La sélection appropriée

des fonctionnalités du noyau est cruciale pour le succès du modèle MVS. Le logiciel STATISTICA offre une large gamme de fonctions de noyau pour les modèles MVS. Pour le noyau gaussien à base radiale (RBF), les paramètres ont été réglés comme suit : perte de fonction ( $\nu$ ) = 1,0000, paramètre de planéité (C) = 10,0000 et paramètre du noyau (Gamma) = 13,93. Ce processus a permis d'identifier les valeurs optimales pour les paramètres cibles du modèle MVS

#### **III.4. Analyse de sensibilité par la méthode de "poids"**

Une analyse de sensibilité est la méthode permettant d'étudier le comportement d'un modèle et d'évaluer l'importance de chaque paramètre d'entrée sur les valeurs de la variable de sortie du modèle. Le but principal de l'analyse de sensibilité est de déterminer comment les données de sortie sont influencées par la variabilité des données d'entrée. À l'aide de ce type d'analyse, il est possible d'évaluer quels paramètres d'entrée doivent être considérés comme les plus significatifs et les moins significatifs pour la prédiction de l'irradiation solaire. Dans notre étude, pour déterminer le paramètre d'entrée le plus influent, une analyse de corrélation est réalisée et des graphiques de régression sont développés en utilisant un réseau de neurones. Cette approche, appelée méthode des "poids", a été initialement introduite par [108] et adoptée par la suite par [109] Elle permet de mesurer l'importance relative (IR) des diverses variables d'entrée sur la sortie [110]. Le processus de calcul se décompose en quatre étapes, comme illustré dans le schéma ci-dessous (Figure III.2). L'objectif de l'analyse de sensibilité est de réduire le nombre de paramètres d'entrée nécessaires pour une prédiction facile de l'irradiation solaire sans augmenter l'erreur du réseau.

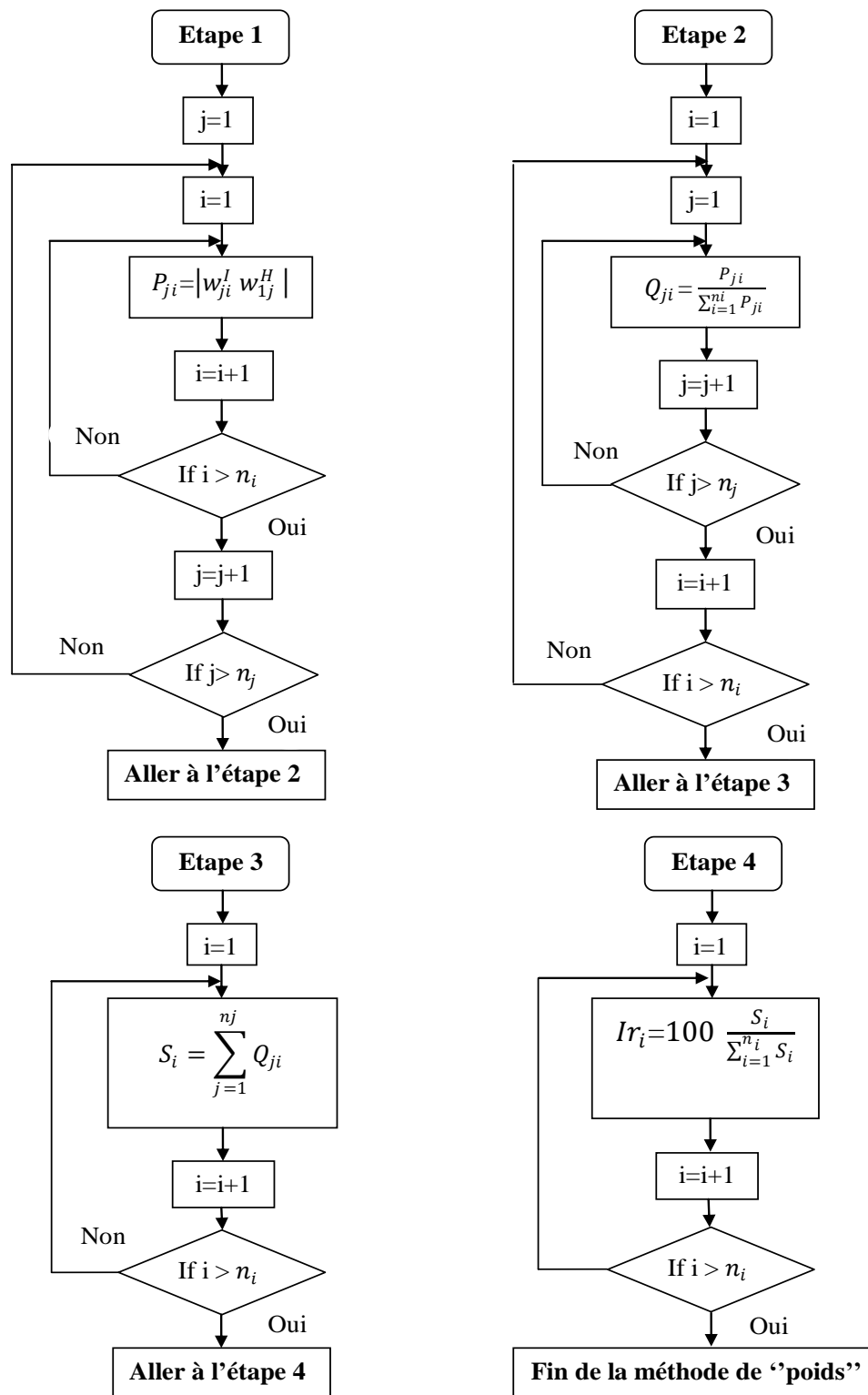


Figure III.2. Diagramme alternatif de la méthode de poids [111].

### III.5 Diagnostic des points aberrants et domaine d'applicabilité du modèle neuronal optimisé

Les valeurs aberrantes sont largement reconnues comme des données distinctes de la tendance générale observée dans un ensemble de données principal. La détection des valeurs aberrantes, qui revêt une importance cruciale dans l'élaboration de modèles, sert à identifier des données individuelles (ou des groupes de données) qui peuvent diverger de l'ensemble de données majoritaire. Dans cette optique, il est essentiel de développer des méthodes robustes pour repérer les valeurs potentiellement suspectes, en vue de filtrer des données expérimentales incorrectes et d'améliorer la précision des modèle [112]. Dans cette étude, la méthode mathématique "Leverage" a été employée pour repérer les valeurs aberrantes. Cette technique commence par le calcul des résidus, puis elle crée une matrice H de points de données d'entrée selon la formule suivante :

$$H = X(X^tX)^{-1}X^t \quad (\text{III.1})$$

Où X désigne une matrice de dimensions  $m \times n$ , où m représente le nombre d'échantillons et n correspond aux paramètres (variables d'entrée) du modèle.

Les valeurs de Hat sont extraites de la diagonale principale de la matrice H.

$$\text{Hat} = \text{diagonal}(H) \quad (\text{III.2})$$

Le diagramme de William, qui illustre les résidus normalisés par rapport au levier (Hat), permet une détection visuelle des valeurs aberrantes. Sur ce graphique, le levier critique ( $H^*$ ) ou seuil est généralement déterminé en utilisant la valeur fournie par l'équation suivante:

$$H^* = \frac{3(n+1)}{m} \quad (\text{III.3})$$

Où m représente le nombre d'échantillons et "n" correspond aux paramètres (variables d'entrée) du modèle. Les résidus normalisés sont calculés en comparant les données de sortie expérimentale avec celles calculées par le modèle.

$$(R\_Norm)_i = \frac{(ISG_i^{\text{exp}} - ISG_i^{\text{cal}})}{\sqrt{\text{Var}(ISG^{\text{exp}} - ISG^{\text{cal}})}} \quad i = 1, \dots, m \quad (\text{III.4})$$

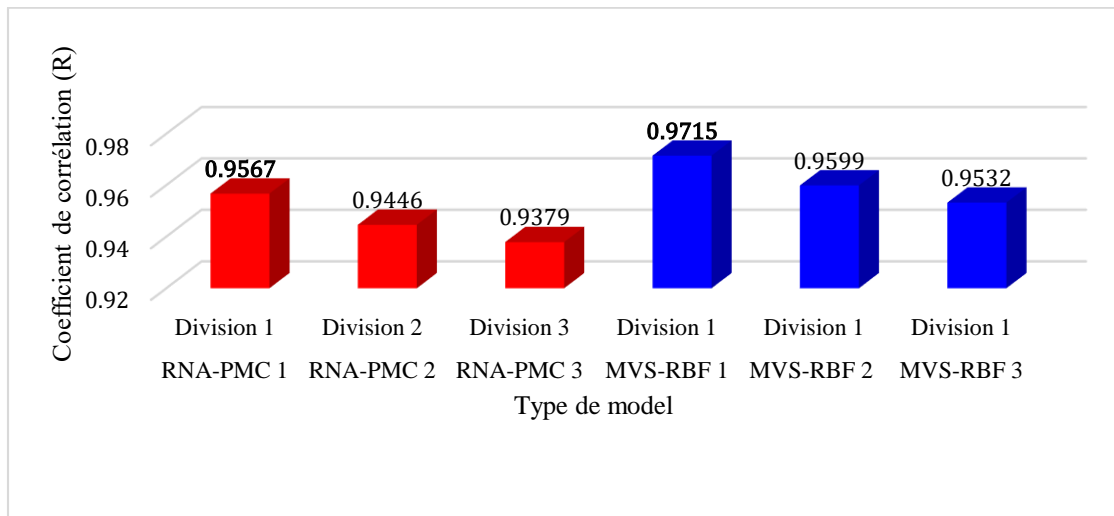
Un résidu normalisé égal à trois est généralement utilisé comme seuil pour déterminer l'acceptation des points de données dans un écart-type qui se situe dans une plage de  $\pm 3$  par rapport à la moyenne.

Cette plage couvre normalement 99% des données dans une distribution normale. Si la plupart des points de données se situent dans les plages de  $0 \leq \text{Hat} \leq H^*$  et de  $-3 \leq R\_Norm \leq 3$ , cela signifie que le modèle a été développé et ses prédictions sont fiables dans son domaine d'application, ce qui en fait un modèle statistiquement valide. Ainsi, on identifie des points de "bon levier élevé" dans la plage de  $0 \leq \text{Hat} \leq H^*$  et de  $-3 \leq R\_Norm \leq 3$ . Les points situés en dehors de cette plage, à  $R\_Norm > 3$  (qu'ils soient supérieurs ou inférieurs à la valeur  $H^*$ ), sont considérés comme des valeurs aberrantes du modèle ou des points de "mauvais levier élevé".

### **III.6. Résultats et discussion**

Cette sous-section présente les résultats des modèles développés dans l'étude pour prédire l'irradiation horizontale globale horaire. Initialement, les données recueillies à la station de Bouzaréah ont généré deux modèles : les réseaux de neurones artificiels de type perceptron à multicouches (RNA-PMC) et la machine à vecteur de support (MVS). Les résultats de ces modèles ont été évalués à l'aide de trois divisions de données différentes pour la phase d'apprentissage, la phase de validation et la phase de test. Les résultats sont présentés dans la Figure III.3, qui illustre les valeurs du coefficient de corrélation (R) obtenues pour chaque division. On peut observer que la division 3 surpasse les deux autres divisions en termes de valeurs de R pour la phase de test, avec  $R = 0,9567$  pour le modèle RNA-PMC et  $R = 0,9715$  pour le modèle SVM. Pour le modèle RNA-PMC, la division 3 comprenait 60 % des données pour la phase d'apprentissage, 20 % pour la validation et 20 % pour la phase de test. Pour le modèle SVM, la division 3 disposait de 60 % des données pour la phase d'apprentissage et de 40 % pour le test. Les résultats suggèrent que la division 3 fournit les prédictions les plus précises, ce qui en fait le choix optimal pour tester les modèles RNA-PMC et SVM.





**Figure III.3.** Effet de la division de la base de données en corrélation des coefficients (R) pour la phase de test.

**Tableau III.4.** Évaluation statistique des erreurs du modèle (RNA-PMC) sur la prédiction de l'irradiation horizontale globale horaire.

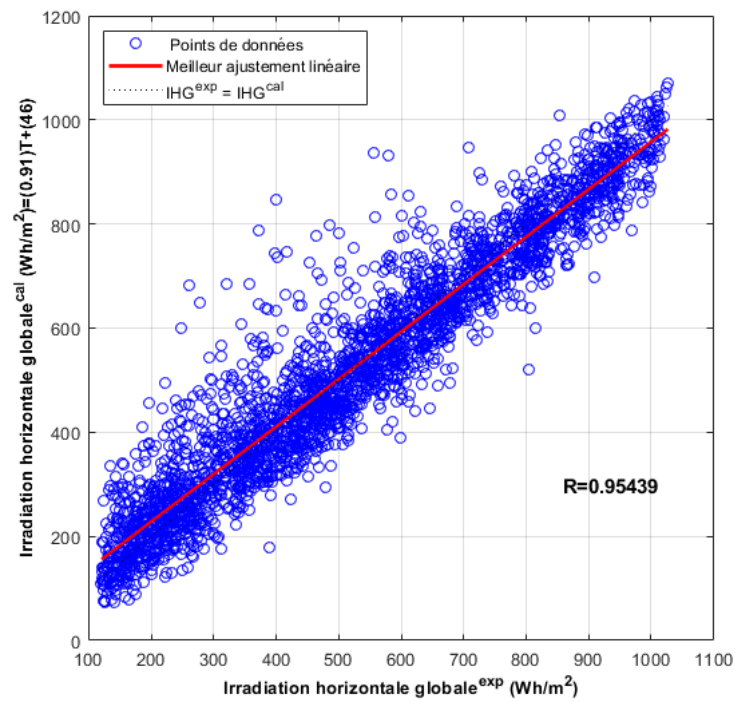
Statistique	Phase de l'apprentissage	Phase de validation	Phase de Test	Totale
<b>R</b>	0.9543	0.9362	0.9567	0.9528
<b>EPM [%]</b>	13.6706	17.3840	14.1855	14.0925
<b>REQM [Wh/m<sup>2</sup>]</b>	72.9729	86.9887	71.4774	74.3451

Les statistiques des erreurs du modèle optimal de RNA-PMC pour la phase d'apprentissage, la phase de validation, la phase de test et la phase totale en termes de R (coefficient de corrélation), Erreur de Prédiction du Modèle (EPM) et la racine de l'Erreur Quadratique Moyenne (REQM) sont présentées dans le Tableau III.4.

La Figure III.4 montre les courbes de régression linéaire de l'irradiation horizontale globale calculée par le modèle neuronal optimal RNA-PMC, optimisé avec les valeurs expérimentales (mesurées) de l'irradiation horizontale globale. Les valeurs de coefficient de corrélation se rapprochent des valeurs idéales (R = 0,9543 avec une REQM de 72,9729 Wh/m<sup>2</sup> en phase d'apprentissage ; R = 0,9362 et REQM = 86,9887 Wh/m<sup>2</sup> en phase de validation ; R = 0,9567 et REQM = 71,4774 Wh/m<sup>2</sup> en phase de test). Le coefficient de corrélation "R" en phase de test mesure la capacité du modèle à interpoler les données expérimentales avec précision. La valeur de corrélation de 0,9567 en phase de test démontre un niveau

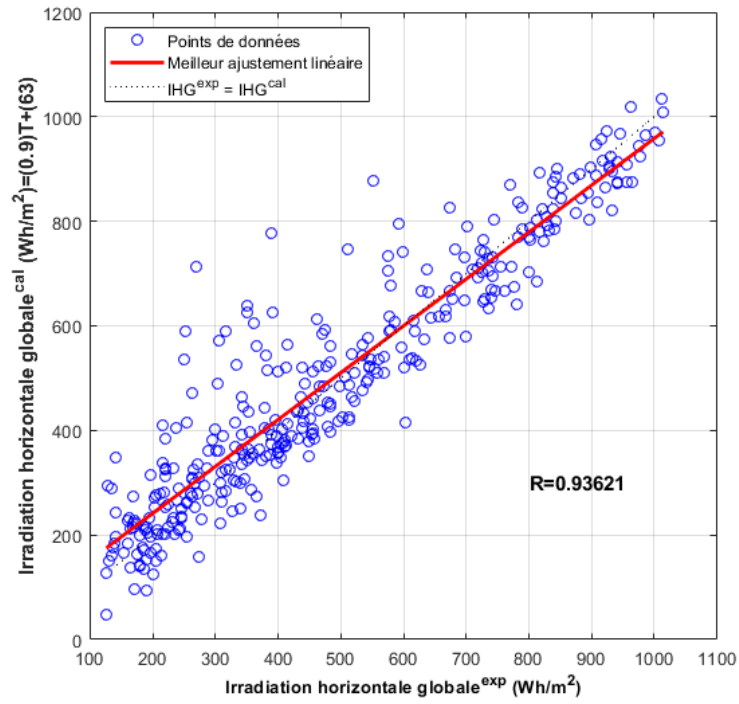
### Chapitre III : Évaluation de l'efficacité de l'apprentissage amélioré dans la prédiction de l'irradiation horizontale globale

élevé de cohérence entre les valeurs expérimentales et les valeurs prédites de l'irradiation horizontale globale, avec  $R = 0,9528$  et  $REQM = 74,3451 \text{ Wh/m}^2$  pour la phase totale.

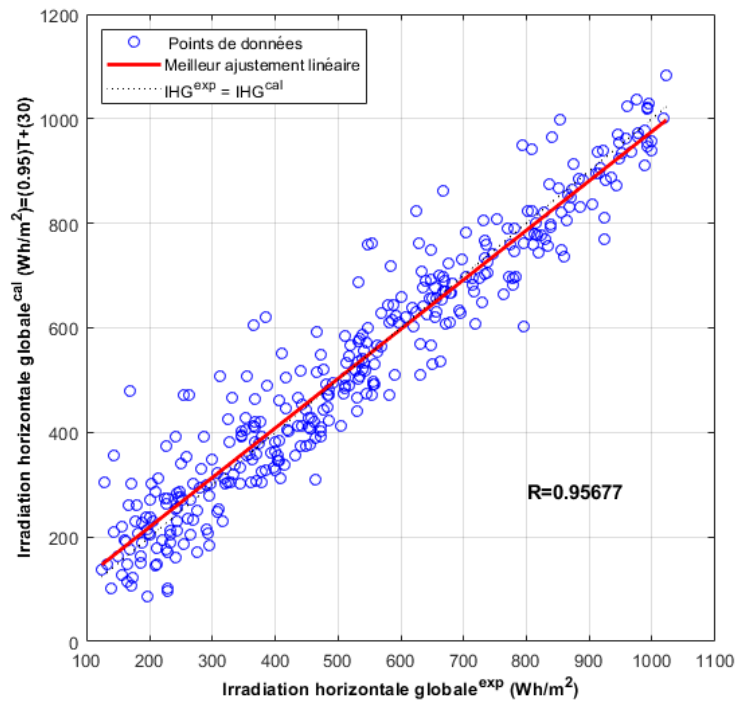


(a)

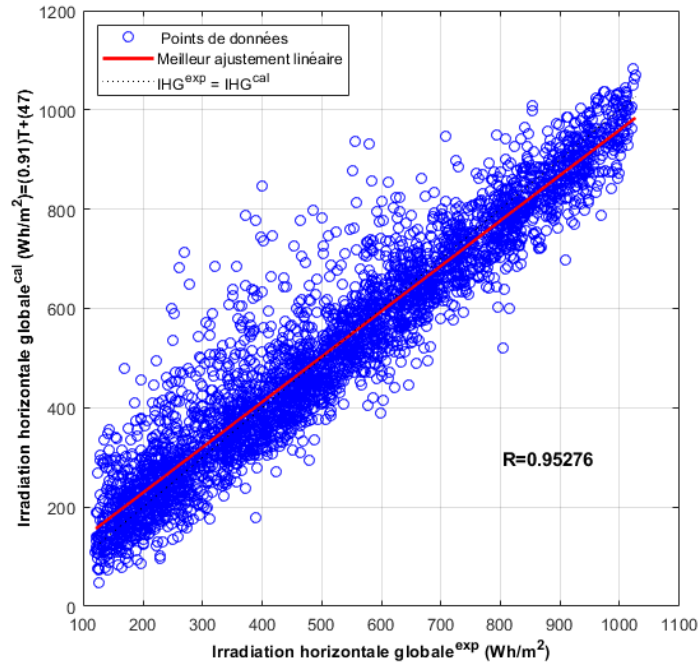
### Chapitre III : Évaluation de l'efficacité de l'apprentissage amélioré dans la prédiction de l'irradiation horizontale globale



(b)



(c)



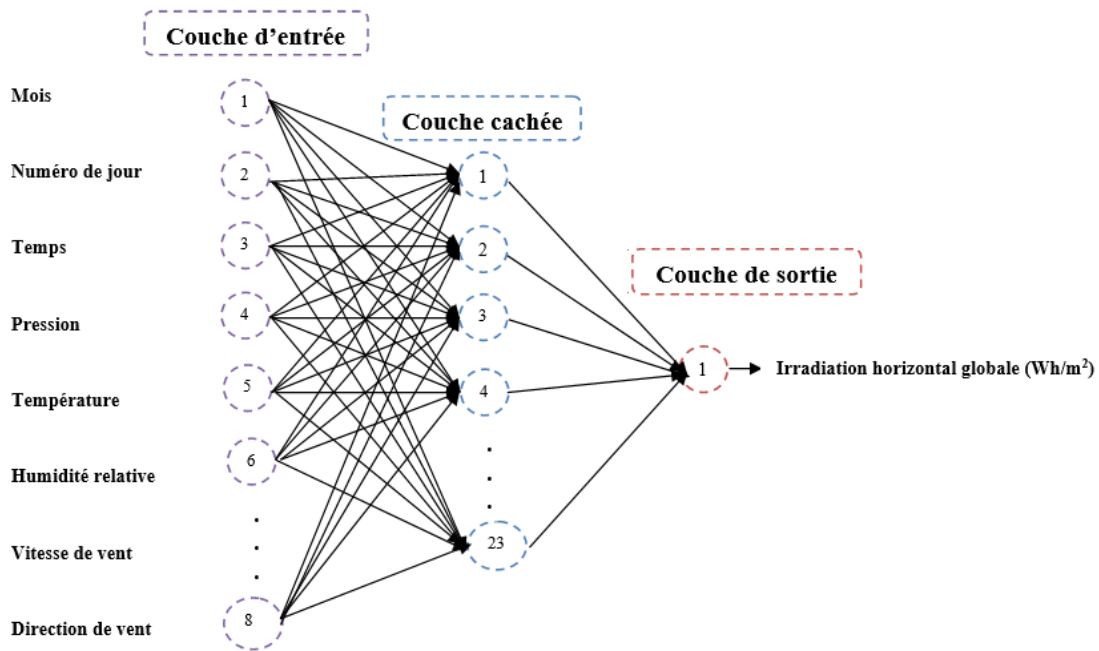
(d)

**Figure III.4.** Modèle les valeurs expérimentales vs. Les valeurs calculées pour les phases de données : (a) phase d'apprentissage, (b) phase de validation, (c) phase de test, (d) totale.

L'architecture optimale du réseau de neurones (RNA-PMC) avec une seule couche cachée est exposée dans la Figure III.5. Nous avons adopté la fonction tangente hyperbolique ( $\tanh$ ) comme fonction de transfert entre la couche d'entrée et la couche cachée, tandis qu'une fonction linéaire (Purelin) a été employée pour relier la couche cachée à la couche de sortie. Dans notre étude, nous avons utilisé l'algorithme d'apprentissage Broyden–Fletcher–Goldfarb–Shanno (BFGS). Les éléments structurels du réseau de neurones utilisés dans la prédiction de l'irradiation solaire globale sont récapitulés dans le Tableau III.5.

**Tableau III.5.** Structure du réseau de neurones optimisé.

Algorithme d'apprentissage	Couche-Entrée	Couche-Cachée		Couche-Sortie	
	Nombre de neurone	Nombre neurone	fonction d'activation	Nombre neurone	fonction d'activation
Broyden–Fletcher–Goldfarb–Shanno quasi-Newton	8	23	tangentes hyperboliques	1	Purelin



**Figure III.5.** Architecture du réseau neuronal.

### III.6.1. Expressions mathématique des modèles RNA-PMC

À partir du RNA-PMC amélioré, tel qu'il est illustré dans la Figure III.6, nous pouvons formuler un modèle mathématique pour représenter l'irradiation solaire en incorporant toutes les variables d'entrée (Mois, Numéro de jour, Temps, Pression, Température, Humidité, Vitesse du vent, Direction du vent) comme suit :

Les sorties  $Z_j$  de la couche cachée :

$$Z_j = f_H \left[ \sum_{i=1}^{08} w_{ji}^I x_i + b_j^H \right] = \frac{\exp(\sum_{i=1}^{08} w_{ji}^I x_i + b_j^H) - \exp(-\sum_{i=1}^{08} w_{ji}^I x_i + b_j^H)}{\exp(\sum_{i=1}^{08} w_{ji}^I x_i + b_j^H) + \exp(-\sum_{i=1}^{08} w_{ji}^I x_i + b_j^H)} \quad (III. 5)$$

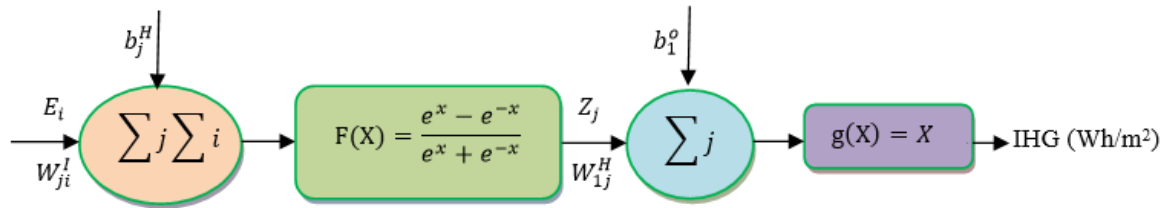
$j=1, 2, \dots, 23$

La sortie " Irradiation horizontale globale "

$$IHG = f_0 \left[ \sum_{j=1}^{23} w_{1j}^H Z_j + b_1^0 \right] = \sum_{j=1}^{23} w_{1j}^H Z_j + b_1^0 \quad (III. 6)$$

En combinant les équations (III.5) et (III.6), nous obtenons la formule mathématique suivante, qui représente l'irradiation solaire globale en prenant en compte toutes les variables d'entrée :

$$IHG = \sum_{j=1}^{23} w_{1j}^H \frac{\exp(\sum_{i=1}^{08} w_{ji}^I x_i + b_j^H) - \exp(-\sum_{i=1}^{08} w_{ji}^I x_i + b_j^H)}{\exp(\sum_{i=1}^{08} w_{ji}^I x_i + b_j^H) + \exp(-\sum_{i=1}^{08} w_{ji}^I x_i + b_j^H)} + b_1^0 \quad (III. 7)$$



**Figure III.6.** Représentation schématique du RNA-PMC amélioré.

Le Tableau III.6 présente les vecteurs de régression linéaire et les performances de plusieurs modèles de Machine à Vecteur de Support (MVS) utilisant diverses fonctions de noyau, notamment les fonctions de noyau linéaires, polynomiales, à base radiale (RBF) et sigmoïdes. Les erreurs de chaque modèle sont évaluées à l'aide de la racine de REQM et R à différentes phases (d'apprentissage, test et totale). Ce tableau met en évidence que les paramètres de régression les plus performants sont ceux qui ont été obtenus grâce au modèle MVS-RBF. Il est évident que les résultats de cette étude démontrent que le modèle MVS-RBF proposé possède une forte capacité à prédire les valeurs de l'irradiation horizontale globale.

**Tableau III.6.** Évaluation des performances des modèles MVS avec différents Fonction noyaux.

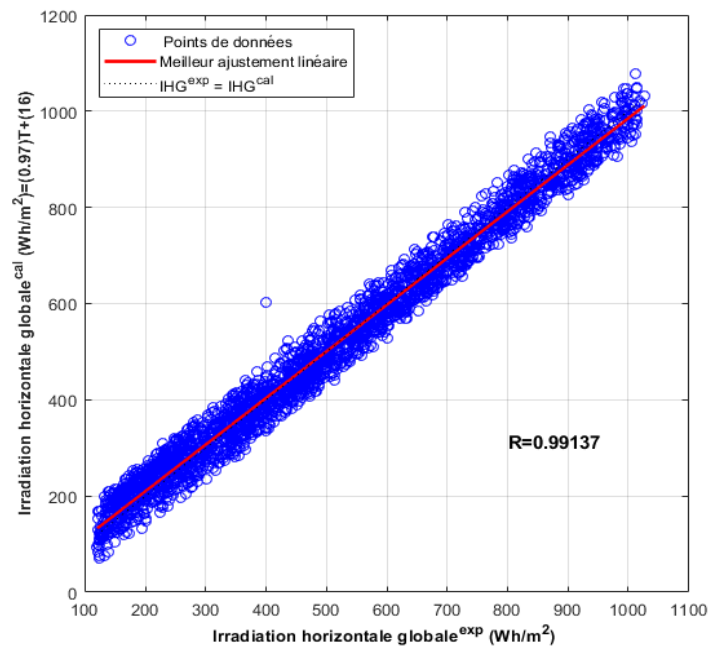
<b>Fonction noyaux</b>	<b>Numéro de SVM</b>	<b>Phase</b>	<b>REQM (Wh/m<sup>2</sup>)</b>	<b>(R)</b>
<b>Linéaire</b>	362	Apprentissage	193.312	0.625
		Test	190.394	0.628
		Totale	192.150	0.626
<b>Polynomial</b>	474	l'apprentissage	118.351	0.884
		Test	120.717	0.873
		Totale	119.303	0.880
<b>Fonctions de base radiales (RBF)</b>	1163	l'apprentissage	32.414	0.991
		Test	57.326	0.972
		Totale	38.704	0.988
<b>Sigmoid</b>	220	l'apprentissage	229.925	0.414
		Test	223.432	0.434
		Totale	227.348	0.422

Le modèle linéaire MVS avec  $C=10$  et  $E=0,1$  a obtenu un REQM de  $193,312 \text{ Wh/m}^2$  et un R de  $0,625$  pour la phase de l'apprentissage, un REQM de  $190,394 \text{ Wh/m}^2$  et un R de  $0,628$  pour la phase de test, et un REQM de  $192,150 \text{ Wh/m}^2$  et un R de  $0,626$  pour la phase totale. Le modèle MVS polynomial avec  $C=10$ ,  $\nu=1$ ,  $\text{degré}=3$  et  $\text{gamma}=0,125$  a obtenu le REQM le plus bas de  $118,351 \text{ Wh/m}^2$  et le R le plus élevé de  $0,884$  dans la phase de l'apprentissage, un REQM de  $120,717 \text{ Wh/m}^2$  et un R de  $0,873$  dans la phase de test, et un REQM de  $119,304 \text{ Wh/m}^2$  et R de  $0,880$  pour la phase totale.

Le modèle MVS-RBF avec  $C=10$ ,  $\nu = 1$  et  $\text{Gamma} = 13,93$  a obtenu la REQM la plus faible de  $32,414 \text{ Wh/m}^2$  et la R la plus élevée de  $0,991$  dans la phase de l'apprentissage, mais la REQM était plus élevée de  $57,326 \text{ Wh/m}^2$  et la R inférieure de  $0,972$  dans la phase de test, et un REQM totale de  $38,706 \text{ Wh/m}^2$  et R de  $0,988$  pour la phase totale.

Le modèle sigmoïde SVM avec  $C=10$ ,  $\nu=0,1$  et  $\text{Gamma}=0,125$  avait le REQM le plus élevé de  $229,925 \text{ Wh/m}^2$  et le R le plus bas de  $0,414$  dans la phase de l'apprentissage, un REQM de  $223,432 \text{ Wh/m}^2$  et un R de  $0,434$  dans la phase de test, et un REQM de  $227,348 \text{ Wh/m}^2$  et un R de  $0,422$  pour la phase totale.

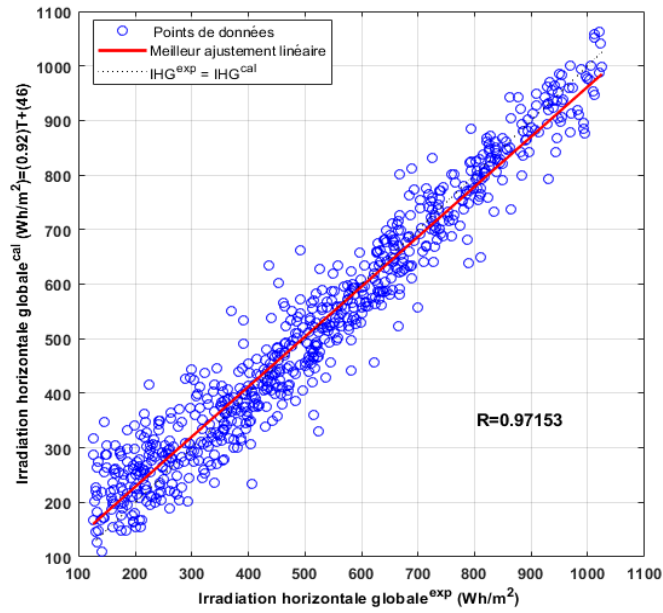
Dans l'ensemble, le modèle MVS-RBF avec  $C=10$ ,  $\nu=1$  et  $\text{Gamma}=13,93$  a surpassé les autres modèles évalués, obtenant le REQM le plus bas et la valeur R la plus élevée durant la phase de l'apprentissage, ainsi que le deuxième REQM le plus bas et le deuxième R le plus élevé durant la phase de test. Il en est résulté le REQM totale le plus bas et la valeur R totale la plus élevée. En comparaison avec les modèles MVS linéaires et sigmoïdes, le modèle MVS-RBF a montré une amélioration substantielle des valeurs REQM et R pendant les phases de l'apprentissage et de test, indiquant son efficacité dans la prédiction de l'irradiation horizontale horaire. En outre, Figure III.7 révèle un alignement robuste entre les valeurs calculées et les valeurs expérimentales recueillies à partir de la station de mesure de l'irradiation solaire, Le coefficient de corrélation sont généralement considéré comme excellent satisfaisant ( $0.9 \leq R < 1.0$ ) pour tous les phases, ces résultats confirme la fiabilité et la précision des modèles MVS-RBF dans la prédiction de l'irradiation horizontal horaire à l'aide des caractéristiques d'entrée choisies.



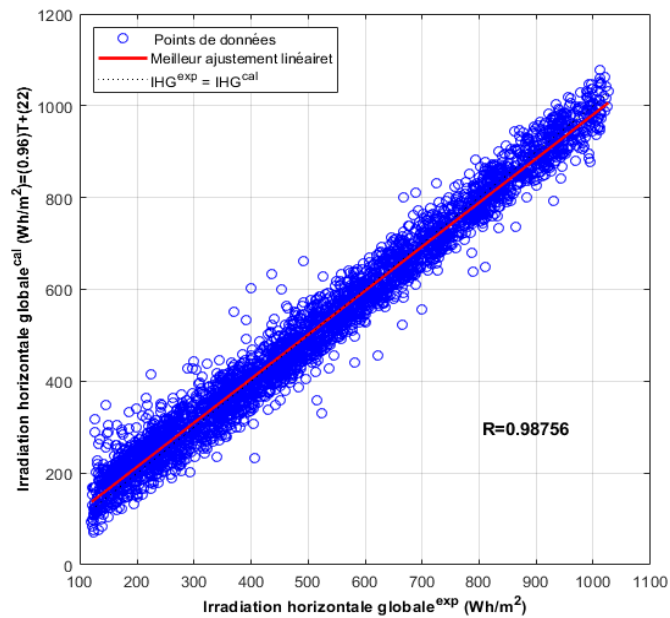
(a)



### Chapitre III : Évaluation de l'efficacité de l'apprentissage amélioré dans la prédiction de l'irradiation horizontale globale



(b)



(c)

**Figure III.7.** Modèle MVS-RBF les valeurs expérimentales vs. Les valeurs calculées pour les phases de données : (a) phase d'apprentissage, (b) phase de test, (c) totale.

**Tableau III.7.** Évaluation les erreurs des modèles élaborés de prédiction de l'irradiation horizontal globale horaire pour les stations Bouzaréah et Tamanrasset.

Model	Phase	Station	R (-)	EMAP (%)	REQM (Wh/m <sup>2</sup> )	Station	R (-)	EAMP (%)	REQM (Wh/m <sup>2</sup> )
<b>RNA-PMC</b>	l'apprentissage	Bouzaréah	0.9544	13.6706	72.9729	Tamanrasset	0.9322	14.7679	109.1829
	Test		0.9568	14.1855	71.4774		0.9520	10.6584	87.2197
	Totale		0.9528	14.0925	74.3451		0.9313	14.9170	109.5552
<b>MVS-RBF</b>	l'apprentissage		0.9914	6.9987	32.4142		0.9204	17.7120	119.0809
	Test		0.9715	12.7947	57.3256		0.9351	13.8780	104.7578
	Totale		0.9876	8.1586	38.7045		0.9231	16.9452	116.3575

Après un processus de vérification complet, nous présentons l'évaluation de la performance des modèles RNA-PMC et MVS-RBF pour la prédiction de l'irradiation horizontale globale aux stations Bouzaréah et Tamanrasset dans le Tableau III.7. Les critères d'évaluation comprennent le coefficient de corrélation (R), l'erreur moyenne absolue en pourcentage (EMAP) et la racine de l'Erreur Quadratique Moyenne (REQM) pour l'ensemble de la phase d'apprentissage, de la phase de test et de la phase totale.

Les deux modèles ont démontré une précision raisonnable dans la prédiction de l'irradiation horizontale globale. Cependant, le modèle MVS-RBF a présenté des performances supérieures par rapport au modèle RNA-PMC en termes de valeurs de R et de REQM pour les deux stations, en particulier pendant la phase de test et la phase d'apprentissage.

Les valeurs de R obtenues par le modèle RNA-PMC sont de 0,9544 et 0,9322 pour les stations Bouzaréah et Tamanrasset, respectivement. En revanche, les valeurs de R obtenues par le modèle MVS-RBF sont sensiblement plus élevées, à 0,9914 et 0,9204 pour les deux stations.

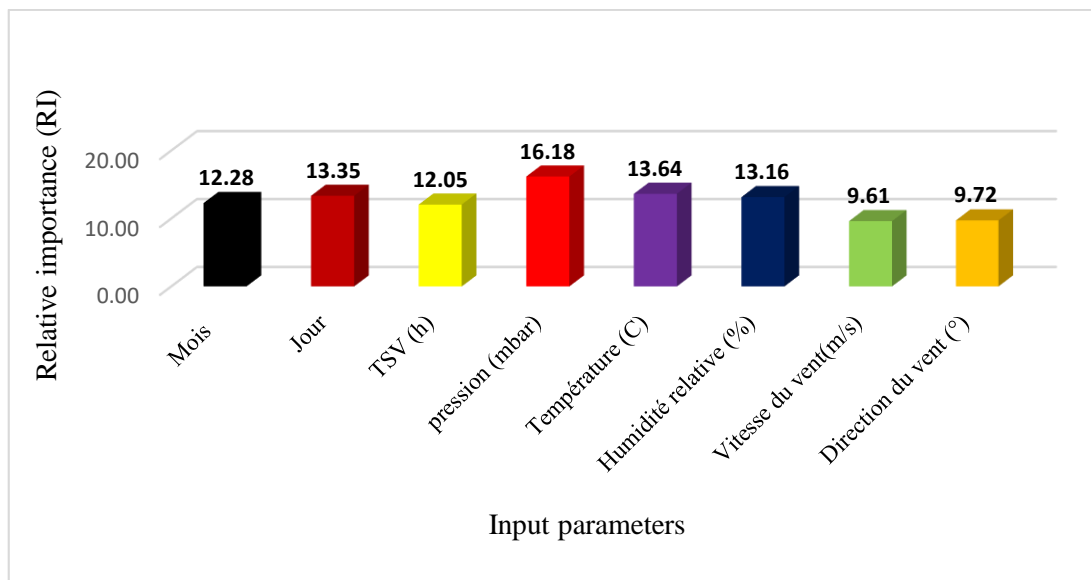
De plus, l'intégration d'inputs supplémentaires tels que le mois, la pression et la direction du vent n'a pas amélioré la performance du modèle. Ces résultats indiquent que le modèle MVS-RBF fournit une prédiction de l'irradiation horizontale globale plus précise que le modèle RNA-PMC.

### **III.6.2. Analyse de la sensibilité**

Pour déterminer l'importance de chaque variable d'entrée sur la prédiction de l'irradiation horizontale globale (Mois, jour, temps solaire vrai (TSV), pression atmosphérique, température, humidité relative,

vitesse du vent et direction du vent) sur la sortie (Irradiation horizontale globale), une analyse de sensibilité a été réalisée en utilisant une approche pondérée, en se basant sur les poids indiqués dans le Tableau III.8. La Figure III.8 illustre l'importance relative des paramètres d'entrée obtenue par cette méthode pour le modèle optimal RNA-PMC.

Les résultats montrent que toutes les variables ont un impact significatif sur l'irradiation horizontale globale. La pression atmosphérique (mbar) présente la contribution la plus élevée (16,18 %). Le mois, le temps solaire vrai (TSV), le jour, l'humidité relative et la température ont presque la même influence sur les valeurs de l'irradiation horizontale globale (13 %). En revanche, la vitesse du vent et la direction du vent ont un effet moindre avec une importance relative de 9,5 %. Ces résultats démontrent que tous les paramètres d'entrée ont une importance relative supérieure à 9 %, ce qui explique l'impact significatif sur la sortie de l'irradiation horizontale globale.



**Figure III.8.** Présentation de l'importance relative (R) des paramètres d'entrées sur la prédiction de l'irradiation horizontale globale.

**Tableau III.8.** Poids et biais du RNA-PMC optimal.

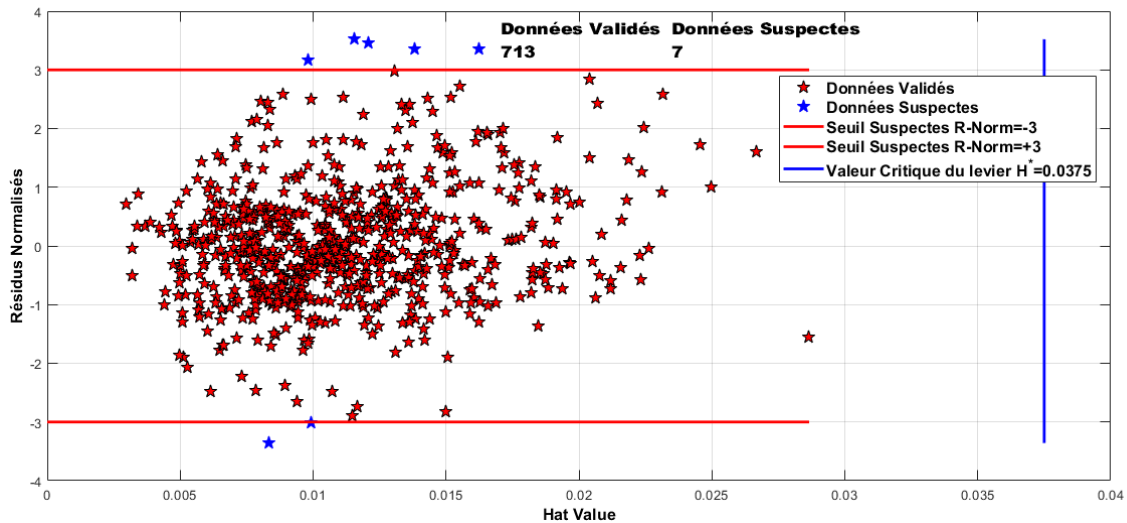
Connexions Couche d'entrée-Couche cachée								Connexions Couche cachée- Couche de sortie		
Poids								Biais	Poids	Biais
$w_{j1}^I$	$w_{j2}^I$	$w_{j3}^I$	$w_{j4}^I$	$w_{j5}^I$	$w_{j6}^I$	$w_{j7}^I$	$w_{j8}^I$	$b_j^H$	$w_{1j}^H$	$w_{1j}^H$
-2.25980	-0.05088	3.92223	-7.45685	-1.70180	1.28522	-1.08103	0.72346	1.42715	0.90212	2.64867
0.14121	-5.46619	7.03822	-2.37513	-0.25117	6.27063	5.38751	2.20857	0.89292	1.37897	
2.24518	-1.37398	0.03130	-1.39437	4.37862	7.35492	-2.04531	-2.44798	-4.49105	-1.29361	
-0.27658	-0.56892	-0.88181	1.61700	-1.47574	-2.34646	-1.46035	1.75833	2.35720	0.70226	
-2.78695	7.98315	-0.33731	-1.61979	1.94647	-1.96786	1.21238	0.04537	-0.41470	-1.25320	
-1.10160	-2.47640	0.12015	2.82219	3.27262	5.72100	-3.06038	-3.48356	-3.19249	-0.50292	
0.91932	-1.97999	1.14696	-3.10859	-3.78266	8.81306	6.92651	1.51372	-4.09118	0.44144	
0.66182	-1.96748	-2.04485	1.53137	4.34448	0.46233	0.98090	-0.97628	2.65306	-0.87961	
8.70700	-5.60228	-2.29323	3.36473	1.98512	0.70634	-1.57209	-0.23767	5.97101	0.32186	
1.06286	7.90311	10.0086	-2.89563	0.91742	4.05550	-1.66980	-2.99851	-4.38294	0.83085	
-5.07550	0.97846	-4.12185	1.40458	-2.62593	4.53764	-1.10454	7.99110	3.90602	0.36236	
-6.27107	-5.15783	1.12022	-8.58965	-2.38857	-2.07397	0.69326	0.74293	-0.39208	-0.86077	
-4.78997	-4.39661	-0.15405	0.69630	-2.65302	-9.13618	3.51204	4.41082	-3.13135	-0.25214	
-1.22046	-2.88745	0.26123	1.84808	4.59306	1.17302	1.21215	3.27029	-1.84627	0.32819	
0.71171	-0.98308	-1.64666	5.09822	-0.46523	-0.12772	-0.08443	-0.19921	-0.87683	0.19692	
4.61649	0.37337	-0.65084	-2.20831	2.06553	5.32560	0.62808	3.86561	-4.18681	0.17612	
0.06448	5.18872	-3.80471	-9.87210	-2.67718	-1.03925	10.40709	1.06140	-4.37227	-1.11621	
0.07592	-1.80472	-0.54358	-8.91482	2.20097	-3.73482	-0.30979	0.80917	0.22551	-1.47291	
4.83923	-1.28543	0.37870	-0.74014	-2.86447	1.29387	4.10436	-0.27869	-4.83809	-1.33203	
1.31037	-0.45636	3.15992	-2.56923	0.43164	-1.04392	-0.85753	-2.85231	-2.21715	0.53609	
3.21823	3.57267	-1.46683	0.09143	2.34572	0.27890	0.41283	-1.09176	3.25463	-3.15938	
-0.52995	-0.96038	5.99382	0.31817	-8.48124	-2.30293	-0.82132	1.69125	-5.19520	0.50673	
-2.90345	0.97881	-2.95479	-0.85923	2.73429	-0.92258	0.43053	-0.74029	-0.93354	0.95469	

### III.6.3. Diagnostic des valeurs aberrantes et domaine applicabilité des modèles optimaux

La figure III.9. Présente le diagramme de William illustrant le domaine d'applicabilité du modèle optimal MSV-RBF (Phase de test). L'existence est représentée dans ce diagramme de 713/720 (99.02%) de points de données valides, Seulement 7/720 (0,98 %) des points de données sont considérés comme suspects. La valeur critique de levier  $H^*$  est calculée comme suit:

$$H^* = \frac{3(n+1)}{m} = \frac{3(8+1)}{720} = 0.0375.$$

Cela démontre que le modèle MVS-RBF optimal est développé et ses prédictions sont effectuées dans un domaine d'applicabilité spécifique, garantissant ainsi la validité statistique du modèle MVS-RBF optimal. En conséquence, on peut conclure que des points "Good High Leverage" sont présents lors de la phase de test.



**Figure III.9.** Diagrammes de Williams pour le modèle MVS-RBF optimal de prédiction de l'irradiation horizontale globale horaire.

La figure III.10. Présente le diagramme de william illustrant le domaine d'applicabilité du modèle réseau neuronal RNA-PMC optimal (Phase de test). L'existence est représentée dans ce diagramme de 355/359 (98.88%) de points de données valides, Seulement 4/355 (1.12%) des points de données sont considérés comme suspects. La valeur critique de levier  $H^*$  est calculée comme suit:

$$H^* = \frac{3(n+1)}{m} = \frac{3(8+1)}{359} = 0.0752.$$

Cela démontre que le modèle RNA-PMC optimal est développé et ses prédictions sont effectuées dans un domaine d'applicabilité spécifique, garantissant ainsi la validité statistique du modèle RNA-PMC optimal. En conséquence, on peut conclure que des points "Good High Leverage" sont présents lors de la phase de test.

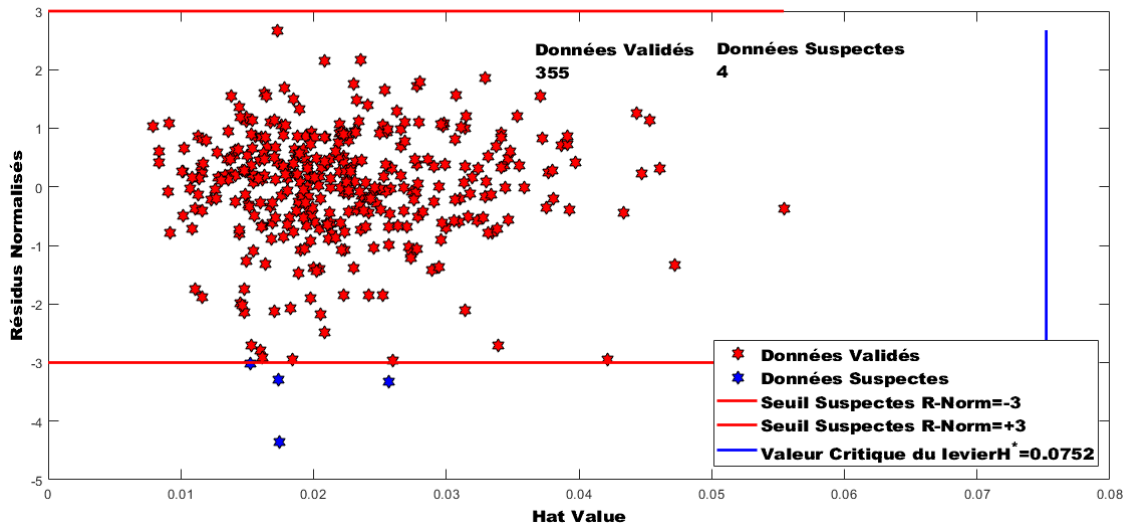


Figure III.10. Diagrammes de Williams pour le modèle RNA-PMC optimal de prédiction de l'irradiation horizontale globale horaire.

### III.7. Conclusion

Cette étude vise à améliorer la précision de la prédiction de l'irradiation horizontale globale horaire en utilisant des techniques avancées d'apprentissage automatique. L'objectif principal est de développer une méthode qui optimise les paramètres des modèles conventionnels d'apprentissage automatique, en particulier les réseaux neuronaux perceptron multicouches RNA-PMC et machine à vecteur de support MVS-RBF.

Pour créer le modèle RNA-PMC le plus efficace, Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-newton a été utilisée comme algorithme de l'apprentissage, quatre fonctions d'activation ont été testées dans la couche cachée et une seule fonction d'activation utilisée dans la couche sortie. En ce qui concerne la puissance et les performances, le modèle MVS avec la fonction noyau de type base radiale (RBF) donne de bien meilleurs résultats que les modèles MVS avec d'autres fonctions. La fonction noyau de type base radiale (RBF) montre également une capacité supérieure pour caractériser les prédictions de l'irradiation horizontale globale du modèle MVS. Les valeurs de différence d'erreur statistique entre le modèle SVM-RBF et le modèle RNA-PMC sont significatives, ce qui indique la plus grande précision du modèle SVM-RBF proposé dans la prédiction de l'irradiation horizontale globale par rapport au modèle RNA-PMC.

En outre, toutes les méthodes d'apprentissage automatique discutées dans cette étude fournissent des prédictions très précises de l'irradiation horizontale globale à différentes résolutions temporelles. Cependant, nos résultats montrent que le modèle MVS-RBF mieux que le modèle RNA-PMC avec fonction d'activation BFGS pour prédire l'irradiation horizontale globale, avec une valeur R de 0,99 et une REQM de 38,70 Wh/m<sup>2</sup> pour la phase totale. En outre, cette étude étudie également les performances des modèles proposés dans différentes régions climatiques de l'Algérie, ce qui est crucial pour prédire avec précision l'irradiation horizontale globale à un endroit spécifique. De cette façon, il pourrait aider dans la conception et l'installation de systèmes d'énergie solaire ainsi que dans l'évaluation des conditions thermiques dans les études de construction.

Cette étude offre une alternative prometteuse aux méthodes traditionnelles actuellement utilisées en Algérie pour prédire l'irradiation horizontale globale. Avec sa précision et ses performances supérieures, le modèle MVS-RBF était un outil précieux pour prédire l'irradiation horizontale globale pour les deux stations, soutenant ainsi le développement et la mise en œuvre des sources d'énergie renouvelables dans le pays. En outre, l'étude ouvre la possibilité d'utiliser ces techniques dans d'autres pays ayant des besoins climatiques et énergétiques similaires.

# Chapitre IV

Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux / agrégation bootstrap



## **IV.1. Introduction**

La quantité totale d'énergie solaire absorbée par la surface terrestre est définie comme l'irradiation solaire global, cette quantité était extrêmement importante dans différents domaines scientifiques tels que l'architecture, l'agriculture, la climatologie et la production d'énergie solaire [113], [114]. Actuellement, l'irradiation solaire et l'énergie renouvelable sont devenues des technologies énergétiques essentielles qui peuvent aider à relever les défis du changement climatique. L'expansion de l'utilisation des sources d'énergie renouvelables entraîne une diminution des émissions de CO<sub>2</sub>, une réduction de la pollution atmosphérique locale, le développement de professions de grande valeur, et une réduction de la dépendance d'un pays aux importations d'énergie fossile. Lorsqu'on compare l'énergie solaire à d'autres sources d'énergie fossiles, il est clair que l'énergie solaire, une source d'énergie importante, n'est pas nocive pour l'environnement et n'a aucun effet sur le réchauffement de la planète [115].

Les données sur l'irradiation solaire sont extrêmement importantes pour les fabricants, les concepteurs de systèmes d'énergie solaire, les architectes et les agriculteurs. Cependant, les mesures de l'irradiation solaire n'ont été prises que dans quelques sites à travers le monde pour diverses raisons dont les coûts d'installation, d'entretien et d'étalonnage. En raison de l'importance des données globales sur l'irradiation solaire plusieurs approches sont utilisées pour prédire l'irradiation solaire à partir de divers paramètres. Plusieurs approches statistiques comme les méthodes d'analyse et les techniques statistiques sont des options possibles pour prédire l'irradiation solaire en utilisant une gamme de facteurs météorologiques [116], [117].

Les résultats de la prédiction de l'irradiation solaire globale horaire ont été publiés dans le "Journal Chemistry in Industry" [103]. Ce chapitre se décompose en quatre axes principaux. Dans le premier axe, nous examinerons la collecte, le prétraitement et l'analyse des données météorologiques. Le deuxième axe sera consacré à l'exploration des techniques les plus récentes pour prédire l'irradiation solaire globale, notamment l'utilisation de réseaux de neurones artificiels (RNA) et bootstrap réseaux de neurones agrégés (BRNA). Le troisième axe abordera l'analyse de la sensibilité, le diagnostic des points aberrants et les domaines d'application. Enfin, le quatrième axe présentera les résultats, la discussion et la comparaison des modèles neuronaux employés.

## **IV.2. Collecte des données météorologiques, prétraitement et analyse**

Les données utilisées dans cette étude ont été collectées à partir de la station radiométrique «Shems », localisée au Centre de Développement des Énergies Renouvelables (CDER) de Bouzaréah, à Alger. Cette station est positionnée à une latitude de 36,8° nord, une longitude de 3,17° est et à une altitude de 345 mètres, comme illustré sur la figure IV.1. La région est caractérisée par un climat méditerranéen, marqué par des étés chauds et secs ainsi que des hivers froids et humides



**Figure IV.1.** Localisation de la station météorologique CDER à Bouzaréah [69].

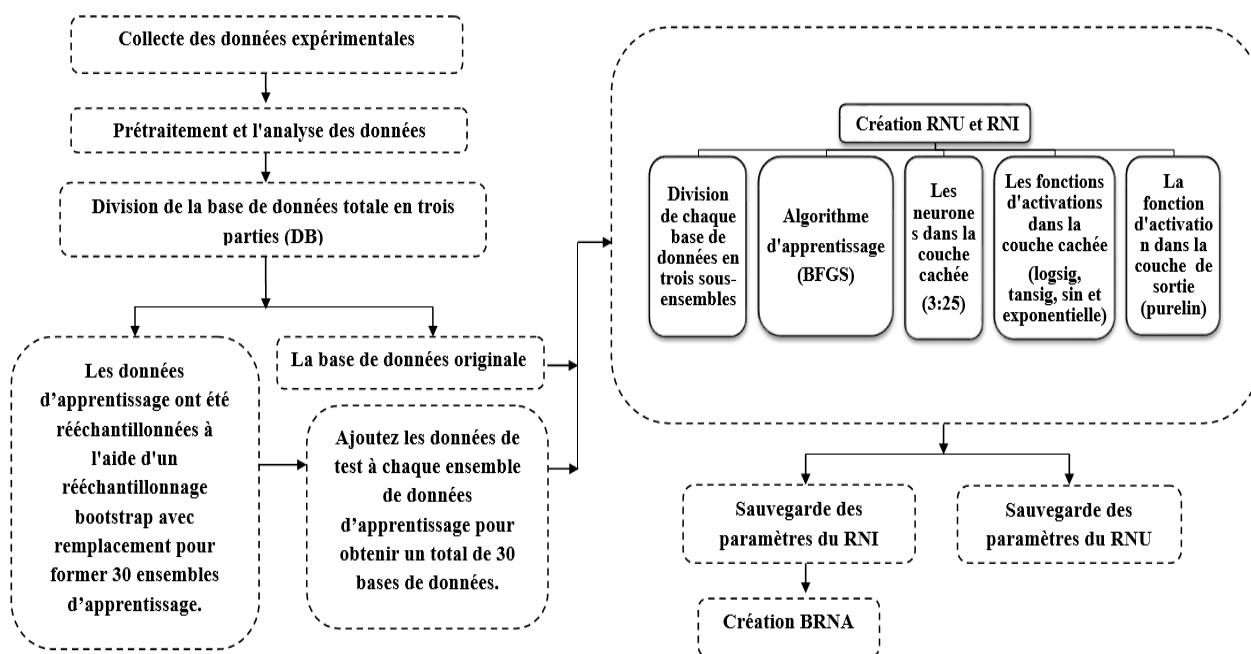
Cette base de données, composée de 3606 points, a été exploitée dans le but d'améliorer les paramètres des bootstrap réseaux de neurones agrégés (BRNA). Au sein de cette base de données, toutes les valeurs inférieures à 120 W/m<sup>2</sup> ont été exclues pour les plages horaires de 5h à 17h, conformément aux directives de l'Organisation météorologique mondiale (WMO) qui définissent la période d'ensoleillement lorsque les valeurs de l'irradiation solaire globale sont supérieures à 120 W/m<sup>2</sup> [118].

## **IV.3. Procédure de modélisation**

### **IV.3.1 Modélisation avec des réseaux neuronaux (RNU et RNI)**

Figure IV. 2 montre une technique de conception et d'amélioration de l'architecture des réseaux neuronaux individuels (RNI) et des réseaux neuronaux uniques (RNU).

## Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux / agrégation bootstrap



**Figure IV.2.** Diagramme de flux pour le développement de réseaux neuronaux (RNU RNI et BRNA<sub>(stacking de 30 réseaux)</sub>).

L'analyse statistique des données d'entrée et de sortie, comprenant les valeurs minimales (min), les moyennes arithmétiques, les valeurs maximales (max), ainsi que les écarts-types (DS), est présentée dans le tableau IV.1.

**Tableau VI.1.** Analyse statistique des données d'entrée et de sortie.

	Min	Moyenne	Max	DS
<b>Mois</b>	1.000	6.484	12.000	3.167
<b>Numéro de jour</b>	1.000	15.715	31.000	8.814
<b>Temps (h)</b>	5.000	11.337	17.000	3.020
<b>Température (K)</b>	280.260	294.682	308.540	6.220
<b>Relative humidité (%)</b>	22.880	58.611	95.000	11.401
<b>Pression (mbar)</b>	972.590	996.482	1020.200	6.242
<b>Vitesse de vent (m/s)</b>	0.080	4.292	16.590	2.227
<b>Direction de vent (°)</b>	0.020	172.060	359.780	118.352
<b>Irradiation solaire Globale (Wh/m<sup>2</sup>)</b>	120.010	515.576	1030.723	247.880

### **IV.3.2 Modélisation avec des bootstrap réseaux de neurones agrégés (BRNA)**

Développer divers modèles de réseaux de neurones artificiels (ANN) et les combiner est une technique attrayante pour accroître la robustesse de ce modèle. De nombreux chercheurs universitaires se sont penchés sur la combinaison de différents modèles de réseaux de neurones individuel afin de construire des modèles bootstrap réseaux de neurones agrégés [119], [120]. L'ensemble de données d'apprentissage a été ré-échantillonné à l'aide d'un ré-échantillonnage par bootstrap avec remplacement pour créer 30 ensembles d'apprentissage [121], [122].

Le bootstrap réseaux de neurones agrégés est représentés à la figure IV.3 lorsque de nombreux modèles de réseaux de neurones indépendants sont développés pour modéliser la même relation, l'équation ci-dessous est utilisée pour les combiner :

$$y = \frac{\sum_{i=1}^n y_i}{n} \quad (\text{IV. 1})$$

Où :  $y_i$  représente la sortie de RNU,  $y$  représente la sortie de BARN et  $n$  est le nombre de modèles de RNI.

L'algorithme d'apprentissage utilisé dans ce travail est le BFGS quasi-Newton (trainbfg). Chaque modèle de réseau de neurones (RNU et RNI) comprend trois couches de neurones : une couche d'entrée avec huit neurones, une couche cachée avec un nombre variable de neurones régulés pendant l'apprentissage, et une couche de sortie avec une unité qui génère la valeur de la prédiction globale de l'irradiation solaire. Le nombre de neurones cachés a varié de trois à vingt-cinq dans cette étude. Les fonctions d'activation telles que la sigmoïde logistique (logsig), la tangente hyperbolique (tanh), la fonction sinusoidale et la fonction d'activation exponentielle ont été appliquées dans la couche cachée. La fonction d'activation pure-linéaire (purelin) a été utilisée dans la couche de sortie. La moyenne des sorties des réseaux neuronaux individuels est utilisée pour calculer la sortie du bootstrap réseaux de neurones agrégés (BRNA). Les outils STATISTICA et MATLAB 2020b ont été utilisés pour appliquer plusieurs RN (RNU et trente RNI) à la prédiction de l'irradiation solaire globale horaire.

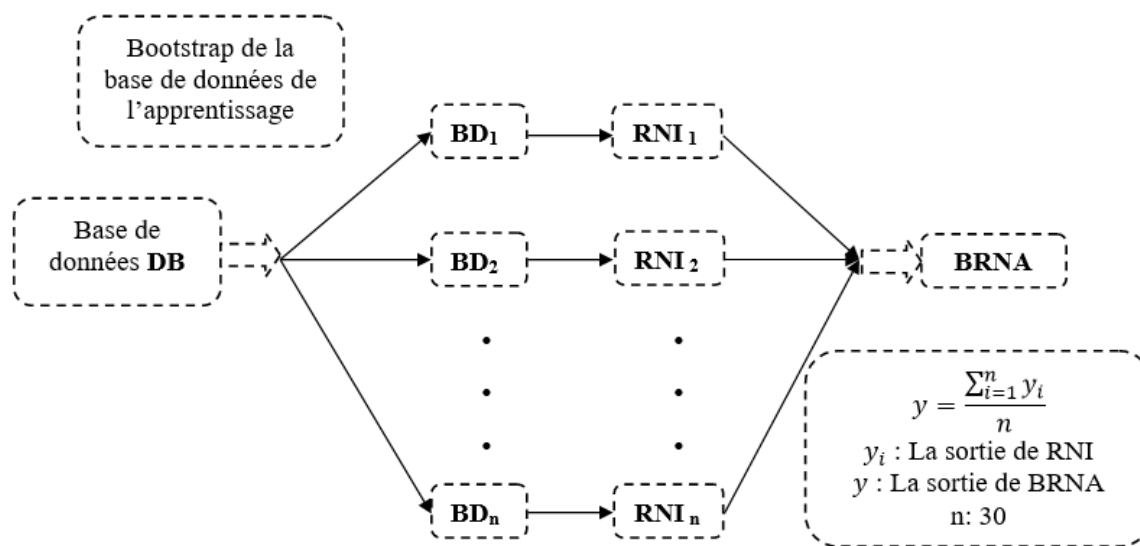


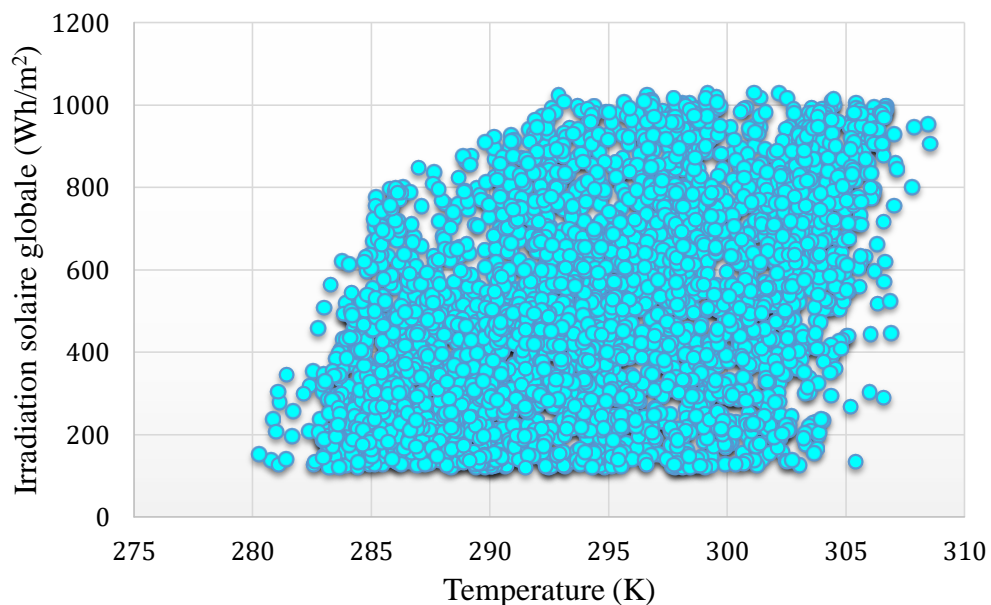
Figure IV.3. Réseaux de neurones agrégés bootstrap.

## IV.4 Résultats et discussion

### IV.4.1 Effet de division de la base des données

Nous avons partitionné la base de données totale en trois ensembles : ensemble 1 (2886 points pour l'apprentissage (80 %), 360 pour la validation (10 %) et 360 pour les tests (10 %)) ; ensemble 2 (2524 points pour l'apprentissage (70 %), 541 pour la validation (15 %) et 541 pour les tests (15 %)) ; et ensemble 3 (2164 points pour l'apprentissage (60 %), 721 pour la validation (20 %) et 721 pour les tests (20 %)).

Figure IV.4 Il montre l'irradiation solaire globale en fonction de la température pour la base de données totale. Il est clair que l'irradiation solaire globale augmente avec l'augmentation de la température à certains endroits.



**Figure IV.4.** Irradiation solaire global Vs température.

Tableau IV.2 montre les coefficients de corrélation ( $R$ ) et les racines de l'erreur quadratique moyenne (REQM) obtenue pour prédire l'irradiation solaire globale horaire sous l'impact de la division de base de données pour le modèle RNU. Les résultats montrent que la première sous ensemble est la division optimale car elle donne les meilleurs résultats que les autres divisions pour la phase de test. Les réseaux neuronaux individuels (RNI) ont ensuite été développés en utilisant la première section de la base de données.

**Tableau IV.2.** Influence de la division de la base de données pour RNU.

	<b>Base de données</b>	<b>REQM (W/m<sup>2</sup>)</b>	<b>R</b>
<b>Division 1</b>	Phase d'apprentissage: 80% "2886 points"	67.2010	0.9624
	Phase de validation : 10% "360 points"	75.3743	0.9540
	phase de test : 10% "360 points"	<b>68.3968</b>	<b>0.9618</b>
<b>Division 2</b>	Phase d'apprentissage: 70% "2524 points"	65.1957	0.9649
	Phase de validation: 15% "541 points"	74.5834	0.9551
	phase de test: 15% "541 points"	73.9391	0.9529
<b>Division 3</b>	Phase d'apprentissage: 60% "2164 points"	69.3676	0.9604
	Phase de validation: 20% "721 points"	75.1032	0.9537
	Phase de test: 20% "721 points"	76.1473	0.9496

#### **IV.4.2 Modèles de performance**

Tableau IV.3 montre les structures des réseaux neuronaux individuels (RNI) et réseaux neuronaux unique (RNU) ne sont pas harmoniques. 15 réseaux neuronaux individuels (RNI) ont appliqué la fonction d'activation log sigmoïde (logistique) et 13 réseaux neuronaux individuels (RNI) ont utilisé la fonction d'activation tangentes hyperboliques (tanh) la même activation qui donne au modèle réseaux neuronaux unique (RNU) son efficacité, sa fiabilité et sa résilience. La fonction d'activation exponentielle a été utilisée dans la couche cachée de deux réseaux neuronaux individuels, Cependant, la fonction sin n'a pas été employée dans la couche cachée des réseaux neuronaux individuels. Ainsi, Nous avons déterminé la suprématie des fonctions d'activation (tanh et sigmoïde) sur les fonctions (sin et exponentielle). Ces résultats corroborent les conclusions de [123] et [124].

**Tableau IV.3.** Structures et performances des modèles RNI.

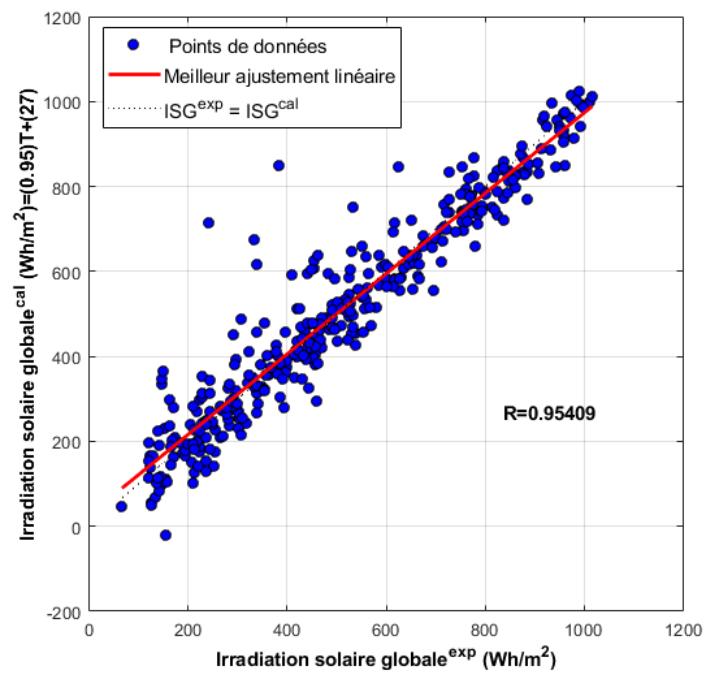
<b>RNI</b>	<b>Nombre de neurones dans la couche cachée</b>	<b>Fonction d'activation dans la couche cachée</b>	<b>RNI</b>	<b>Nombre de neurones dans la couche cachée</b>	<b>Fonction d'activation dans la couche cachée</b>
RNI <sub>1</sub>	22	Tanh	RNI <sub>16</sub>	24	sigmoïde
RNI <sub>2</sub>	22	sigmoïde	RNI <sub>17</sub>	24	Exponential
RNI <sub>3</sub>	23	Tanh	RNI <sub>18</sub>	25	Tanh
RNI <sub>4</sub>	21	sigmoïde	RNI <sub>19</sub>	22	sigmoïde
RNI <sub>5</sub>	23	sigmoïde	RNI <sub>20</sub>	11	Tanh
RNI <sub>6</sub>	20	Exponential	RNI <sub>21</sub>	17	sigmoïde
RNI <sub>7</sub>	21	Tanh	RNI <sub>22</sub>	22	Tanh
RNI <sub>8</sub>	17	sigmoïde	RNI <sub>23</sub>	20	sigmoïde
RNI <sub>9</sub>	19	sigmoïde	RNI <sub>24</sub>	10	Tanh
RNI <sub>10</sub>	19	sigmoïde	RNI <sub>25</sub>	19	Tanh
RNI <sub>11</sub>	14	sigmoïde	RNI <sub>26</sub>	18	sigmoïde
INN <sub>12</sub>	24	Tanh	INN <sub>27</sub>	23	Tanh
INN <sub>13</sub>	22	Tanh	INN <sub>28</sub>	15	Tanh
INN <sub>14</sub>	21	Logsig	INN <sub>29</sub>	21	logsig
INN <sub>15</sub>	20	Logsig	INN <sub>30</sub>	24	Tanh

Selon cette analyse, deux types de modèles de réseaux neuronaux ont été développés (RNU et BRNA<sub>(Stacking of 30 réseaux)</sub>) dans notre étude pour prédire l'irradiation solaire global horaire sur un plan horizontal. Figure IV.5 présente une comparaison entre les valeurs expérimentales et calculées de l'irradiation solaire global horaire avec des vecteurs d'accord proches de l'optimisation des profils de réseaux neuronaux, pour RNU ( $R = 0,9540$  pour la phase de validation et  $R = 0,9618$  pour la phase de

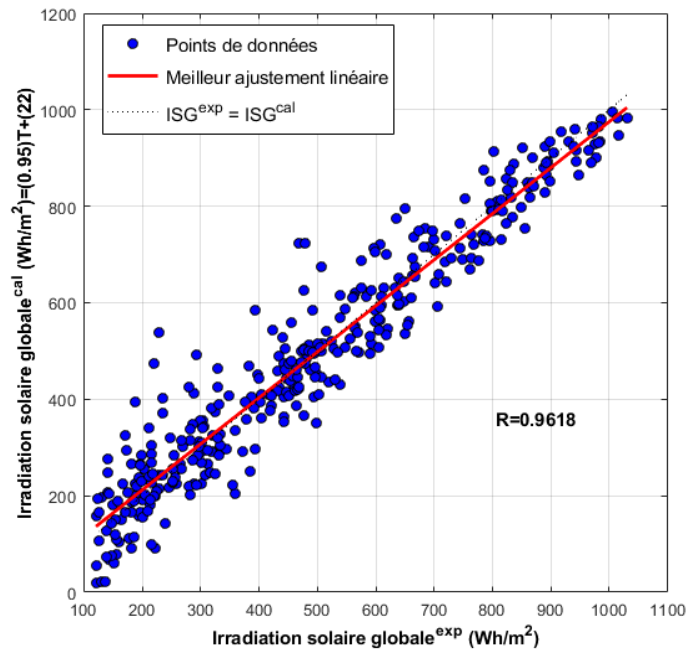


## Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux / agrégation bootstrap

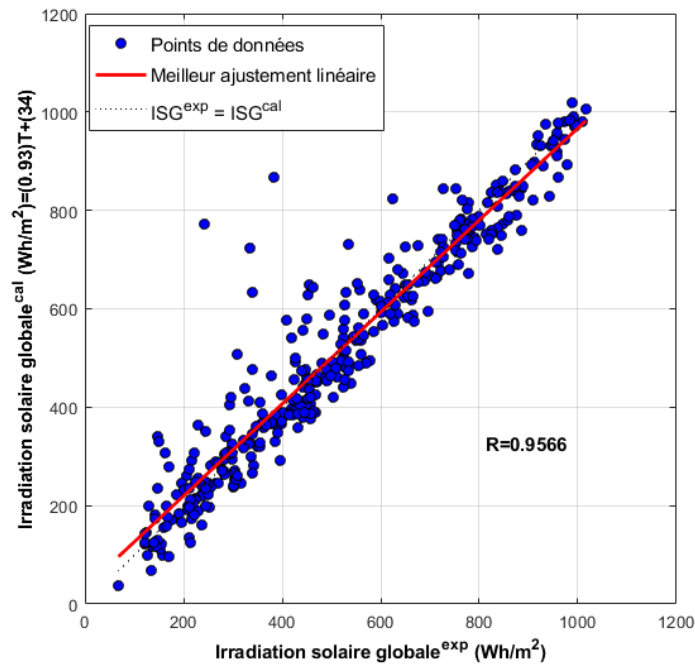
test) et pour BRNA (Stacking of 30 réseaux) ( $R = 0,9566$  pour la phase de validation et  $0,9680$  pour la phase de test). Dans les deux modèles (RNU et BRNA), la pente est proche de 1 pour la phase de validation et elle est extrêmement proche de 1 pour la phase de. L'interception  $b$  est éloignée de 0 pour la phase de validation et de test dans les deux modèles (RNU et BRNA). En général, les coefficients de corrélation sont considérés comme excellents lorsque ( $0,9000 \leq R \leq 1,0000$ ) pour ces modèles (RNU et BRNA) ; cela démontre la résilience des modèles de réseaux de neurones existants ainsi que leur capacité à prédire l'irradiation solaire global horaire.



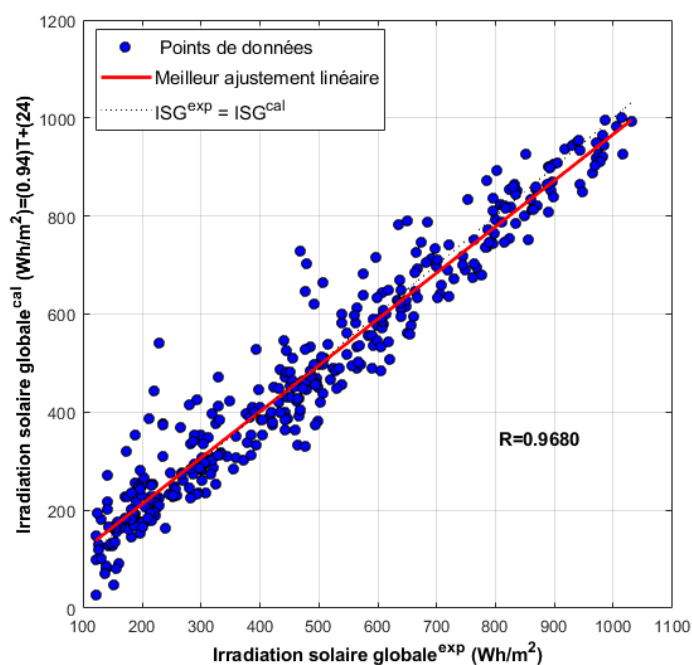
(a)



(b)



(c)



(d)

**Figure IV.5.** Comparaison des valeurs expérimentales Vs. valeurs calculées : (a) la phase de validation (RNU), (b) la phase de test (RNU), (c) la phase de validation (RNAB), (d) la phase de test (RNAB).

#### IV.4.3 Comparaison entre RNI, RNU et BRNA

Tableau IV.4. Montre les erreurs de 30 modèles RNI, RNU et BRNA pour l'apprentissage, la validation, le test et l'ensemble total. Les résultats de ces réseaux sont divergents pour tous les ensembles. Les résultats de ces réseaux ne sont pas uniformes dans tous les ensembles un réseau avec un minimum d'erreurs dans l'ensemble de données d'apprentissage peut avoir de grosses erreurs dans l'ensemble de données de test. Les erreurs REQM et ESP les plus faibles de RNI<sub>13</sub> pour la phase de test étaient 67,5264 Wh/m<sup>2</sup> et 13,8227 % respectivement. De plus les erreurs REQM et ESP du modèle RNAB pour la phase de test sont de 62,4856 Wh/m<sup>2</sup> et de 12,7908 % respectivement. La précision du modèle RNBA est considérablement améliorée en combinant plusieurs modèles RNI. Les deux approches ont été évaluées en terme REQM, ESP, AEM et EMP pour établir le modèle généré BRNA comme plus crédible que les modèles RNU.

**Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux /  
agrégation bootstrap**

**Tableau IV.4.** Performances des modèles RNI, RNU et BNRA en termes de R, EAM, EMP, REQM et ESP.

		<b>R</b>	<b>AEM (W/m<sup>2</sup>)</b>	<b>EMP (%)</b>	<b>REQM (W/m<sup>2</sup>)</b>	<b>ESP (%)</b>
<b>RNI<sub>1</sub></b>	Phase d'apprentissage	0.9610	47.1435	12.5882	67.1635	12.9398
	Phase de validation	0.9525	51.1208	14.4674	76.6738	15.1604
	phase de test	0.9600	52.1814	15.4310	69.8713	14.3027
	phase totale	0.9600	48.0437	13.0597	68.4441	13.2986
<b>RNI<sub>2</sub></b>	Phase d'apprentissage	0.9640	47.5843	12.6655	65.7916	12.5841
	Phase de validation	0.9493	55.2847	16.3270	79.4781	15.7149
	phase de test	0.9607	53.5165	15.9460	69.3198	14.1898
	phase totale	0.9622	48.9456	13.3587	67.6364	13.0651
<b>RNI<sub>3</sub></b>	Phase d'apprentissage	0.9668	45.0790	12.2818	63.6719	12.3069
	Phase de validation	0.9422	52.1937	15.2352	84.6472	16.7369
	phase de test	0.9626	51.9116	15.3468	67.8110	13.8809
	phase totale	0.9638	46.4718	12.8828	66.4765	12.9502
<b>RNI<sub>4</sub></b>	Phase d'apprentissage	0.9572	51.3818	13.7926	72.1796	13.8091
	Phase de validation	0.9479	53.9462	15.4317	80.0534	15.8286
	phase de test	0.9577	54.7978	15.9769	72.4171	14.8238
	phase totale	0.9563	51.9790	14.1744	73.0275	14.1091
<b>RNI<sub>5</sub></b>	Phase d'apprentissage	0.9589	49.1405	13.2518	69.2011	13.5220

**Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux /  
agrégation bootstrap**

---

	Phase de validation	0.9492	54.0600	15.9739	78.8503	15.5907
	phase de test	0.9612	51.8924	15.4010	68.9998	14.1243
	phase totale	0.9581	49.9066	13.7383	70.2045	13.7968
<b>RNI<sub>6</sub></b>	Phase d'apprentissage	0.9583	50.2839	13.9636	71.3760	14.0178
	Phase de validation	0.9452	57.0546	15.8286	82.114	16.2361
	phase de test	0.9579	52.9846	15.3211	71.8980	14.7175
	phase totale	0.9568	51.2297	14.2854	72.5713	14.3202
<b>RNI<sub>7</sub></b>	Phase d'apprentissage	0.9601	50.6206	13.6161	69.9531	13.4522
	Phase de validation	0.9467	54.5593	15.6510	81.2043	16.0562
	phase de test	0.9589	54.7011	15.8990	70.9891	14.5315
	phase totale	0.9586	51.4215	14.0472	71.2591	13.8248
<b>RNI<sub>8</sub></b>	Phase d'apprentissage	0.9560	50.5514	13.5867	71.5767	13.8712
	Phase de validation	0.9434	55.6797	15.7177	83.4062	16.4916
	phase de test	0.9595	54.6889	15.5641	70.8372	14.5004
	phase totale	0.9550	51.4767	13.9970	72.7722	14.2066
<b>RNI<sub>9</sub></b>	Phase d'apprentissage	0.9586	50.1590	13.2991	70.4095	13.5367
	Phase de validation	0.9492	52.2051	14.1481	79.0788	15.6359
	phase de test	0.9579	52.9963	14.8959	71.5429	14.6448
	phase totale	0.9575	50.6467	13.5433	71.4351	13.8563

**Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux /  
agrégation bootstrap**

---

<b>RNI<sub>10</sub></b>	Phase d'apprentissage	0.9605	47.8047	13.1407	68.1869	13.2630
	Phase de validation	0.9476	53.1083	15.2133	80.8801	15.9921
	phase de test	0.9568	55.2127	16.4591	73.2162	14.9874
	phase totale	0.9587	49.0741	13.6791	70.0673	13.7192
<b>RNI<sub>11</sub></b>	Phase d'apprentissage	0.9592	50.8677	13.4755	70.2013	13.5973
	Phase de validation	0.9440	57.5783	16.8159	82.7634	16.3645
	phase de test	0.9558	56.9092	16.8197	73.6745	15.0812
	phase totale	0.9573	52.1412	14.1431	71.9028	14.0309
<b>RNI<sub>12</sub></b>	Phase d'apprentissage	0.9607	48.9734	13.0338	68.1828	13.1644
	Phase de validation	0.9447	55.1454	15.9311	82.1242	16.2381
	phase de test	0.9565	55.9508	16.0230	72.6823	14.8781
	phase totale	0.9586	50.2865	13.6216	70.1531	13.6543
<b>RNI<sub>13</sub></b>	Phase d'apprentissage	0.9654	46.0540	12.0899	65.4828	12.3316
	Phase de validation	0.9486	56.2848	15.7905	79.3915	15.6978
	phase de test	0.9629	50.8105	14.5662	67.5264	13.8227
	phase totale	0.9635	47.5507	12.7067	67.2039	12.8190
<b>RNI<sub>14</sub></b>	Phase d'apprentissage	0.9578	50.4869	13.3521	70.4352	13.4370
	Phase de validation	0.9495	52.2278	14.5857	78.6023	15.5417
	phase de test	0.9571	53.1293	15.1674	72.2901	14.7978

**Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux /  
agrégation bootstrap**

	phase totale	0.9570	50.9246	13.6566	71.4780	13.7779
<b>RNI<sub>15</sub></b>	Phase d'apprentissage	0.9628	47.0259	12.2781	66.6456	12.7380
	Phase de validation	0.9502	52.9697	14.6582	78.4081	15.5033
	phase de test	0.9595	51.8923	15.0680	70.4067	14.4122
	phase totale	0.9612	48.1054	12.7944	68.2898	13.1834
<b>RNI<sub>16</sub></b>	Phase d'apprentissage	0.9652	44.8391	11.8842	63.8780	12.2965
	Phase de validation	0.9490	51.3384	14.7853	79.4129	15.7020
	phase de test	0.9608	51.5260	15.4823	69.8059	14.2893
	phase totale	0.9630	46.1559	12.5332	66.1954	12.8530
<b>RNI<sub>17</sub></b>	Phase d'apprentissage	0.9566	52.3531	14.0740	71.8801	13.9894
	Phase de validation	0.9434	54.9935	15.4795	83.5734	16.5246
	phase de test	0.9560	57.0916	16.4221	73.9104	15.1295
	phase totale	0.9551	53.0900	14.4489	73.3337	14.3655
<b>INN<sub>18</sub></b>	Phase d'apprentissage	0.9690	44.2419	11.7796	61.8780	11.9782
	Phase de validation	0.9432	56.3813	15.8428	84.1506	16.6388
	phase de test	0.9566	54.7638	16.2812	73.0082	14.9448
	phase totale	0.9650	46.5049	12.6349	65.6018	12.7953
<b>RNI<sub>19</sub></b>	Phase d'apprentissage	0.9610	49.6754	13.6400	68.4347	13.3602
	Phase de validation	0.9440	56.4846	15.9788	83.2786	16.4663

**Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux /  
agrégation bootstrap**

---

	phase de test	0.9583	54.9811	16.2663	71.4464	14.6251
	phase totale	0.9589	50.8852	14.1358	70.3582	13.8170
<b>RNI<sub>20</sub></b>	Phase d'apprentissage	0.9608	49.6520	13.1038	69.0310	13.0658
	Phase de validation	0.9417	54.5626	15.7424	84.6982	16.7470
	phase de test	0.9542	55.8594	16.3940	74.5202	15.2543
	phase totale	0.9582	50.7622	13.6958	71.3056	13.6574
<b>RNI<sub>21</sub></b>	Phase d'apprentissage	0.9612	48.7132	13.0784	68.8432	13.1080
	Phase de validation	0.9456	53.5130	14.8187	82.0312	16.2197
	phase de test	0.9582	54.1982	16.0784	71.1230	14.5589
	phase totale	0.9593	49.7402	13.5518	70.4979	13.5679
<b>RNI<sub>22</sub></b>	Phase d'apprentissage	0.9640	47.4297	12.7914	66.6506	12.9129
	Phase de validation	0.9504	51.9073	14.5081	78.0722	15.4369
	phase de test	0.9577	54.3245	16.0705	71.6107	14.6587
	phase totale	0.9620	48.5653	13.2903	68.3802	13.3462
<b>RNI<sub>23</sub></b>	Phase d'apprentissage	0.9610	47.7385	12.9521	67.4997	13.1605
	Phase de validation	0.9474	52.5515	14.9656	80.3129	15.8799
	phase de test	0.9595	53.1870	15.8274	70.5140	14.4342
	phase totale	0.9594	48.7633	13.4403	69.1873	13.5729
<b>RNI<sub>24</sub></b>	Phase d'apprentissage	0.9502	55.3574	14.7447	77.5343	14.9675



**Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux /  
agrégation bootstrap**

---

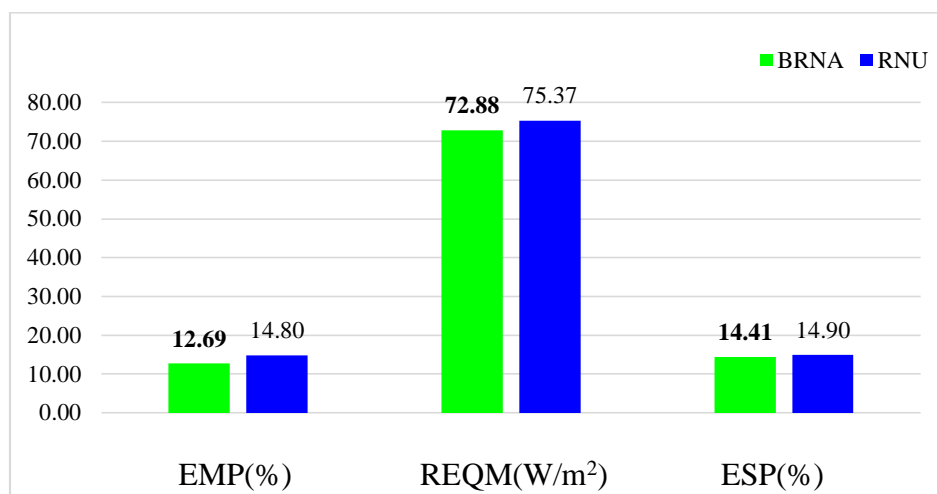
	Phase de validation	0.9430	57.2350	15.7099	83.2881	16.4682
	phase de test	0.9554	54.8173	15.5572	73.8441	15.1159
	phase totale	0.9500	55.4910	14.9222	77.7701	15.1349
<b>RNI<sub>25</sub></b>	Phase d'apprentissage	0.9653	44.9974	12.1271	64.3141	12.3995
	Phase de validation	0.9561	48.8730	13.5462	73.5154	14.5359
	phase de test	0.9610	52.3215	15.2251	69.0457	14.1337
	phase totale	0.9640	46.1158	12.5782	65.7720	12.7867
<b>RNI<sub>26</sub></b>	Phase d'apprentissage	0.9629	47.2317	12.5265	67.0250	12.6602
	Phase de validation	0.9443	54.5213	15.7815	82.9192	16.3953
	phase de test	0.9585	53.2841	16.2463	71.0661	14.5472
	phase totale	0.9606	48.5640	13.2230	69.1814	13.2286
<b>RNI<sub>27</sub></b>	Phase d'apprentissage	0.9633	45.7982	12.1920	65.1769	12.6293
	Phase de validation	0.9566	52.2369	13.7019	71.7125	13.6813
	phase de test	0.9578	54.5804	15.5618	71.5605	14.6484
	phase totale	0.9561	52.4105	13.9793	72.3418	13.9450
<b>RNI<sub>28</sub></b>	Phase d'apprentissage	0.9566	52.2369	13.7019	71.7125	13.6813
	Phase de validation	0.9506	51.6319	14.6198	77.9242	15.4076
	phase de test	0.9578	54.5804	15.5618	71.5605	14.6484
	phase totale	0.9561	52.4105	13.9793	72.3418	13.9450

**Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux /  
agrégation bootstrap**

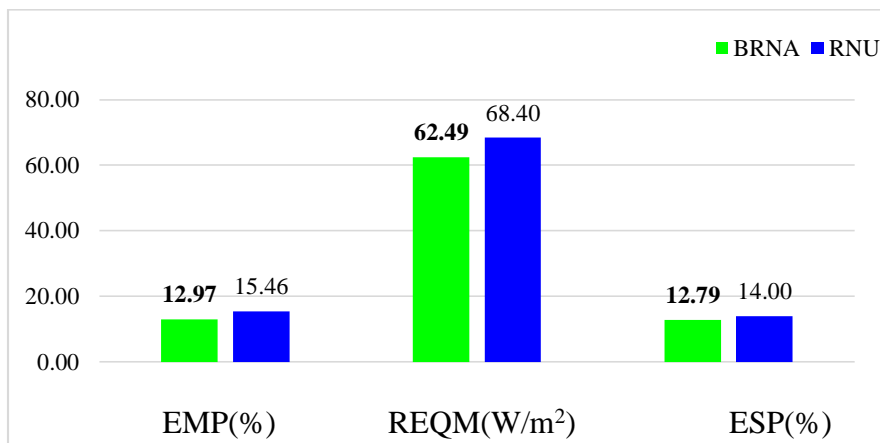
---

<b>RNI<sub>29</sub></b>	Phase d'apprentissage	0.9613	46.8625	12.0862	66.7803	12.7641
	Phase de validation	0.9455	55.7374	15.2728	81.8007	16.1741
	phase de test	0.9612	52.6002	14.6511	68.9905	14.1224
	phase totale	0.9596	48.3217	12.6605	68.6472	13.2527
<b>RNI<sub>30</sub></b>	Phase d'apprentissage	0.9642	45.7545	12.1616	65.0989	12.7690
	Phase de validation	0.9481	51.8555	14.6248	79.7014	15.7590
	phase de test	0.9591	53.3470	15.2213	70.4295	14.4169
	phase totale	0.9620	47.1219	12.7132	67.2396	13.2547
<b>RNU</b>	Phase d'apprentissage	0.9624	47.8535	12.8494	67.2010	12.9266
	Phase de validation	0.9541	50.6691	14.7987	75.3743	14.9034
	phase de test	0.9618	51.3457	15.4556	68.3968	14.0008
	phase totale	0.9615	48.4834	13.3044	68.1802	13.2305
<b>BRNA</b> (stacking 30 réseaux)	Phase d'apprentissage	0.9566	46.0080	12.6926	72.8758	14.4094
	Phase de validation	0.9680	46.3138	12.9661	62.4856	12.7908

Cette comparaison des modèles RNAB et RNU est illustrée par les Figure IV.6 (a) et (b). L'avantage du modèle RNAB par rapport au modèle RNU a été démontré par une comparaison des résultats de la phase de validation et de test. Cela démontre la robustesse des modèles RNU et du modèle RNBA, qui a une plus grande exactitude et peut être plus précis et fiable dans la prédiction de l'irradiation solaire globale lorsque des données invisibles sont appliquées.



(a)



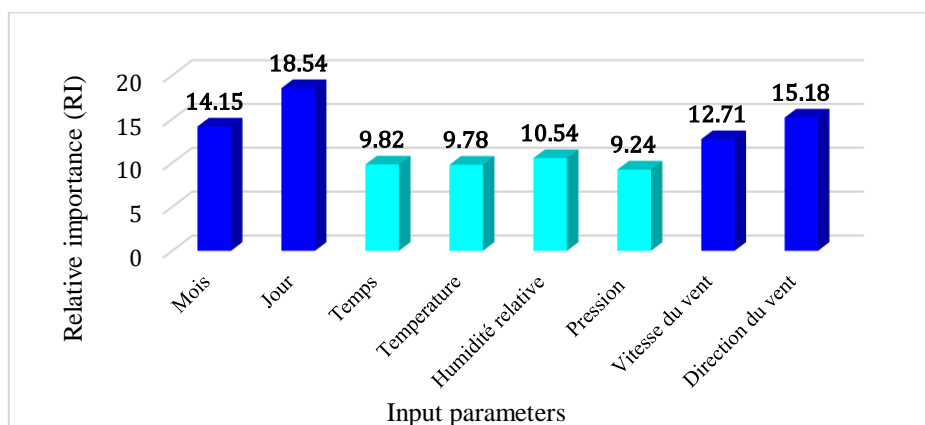
(b)

**Figure IV.6.** EMP, REQM, et ESP pour BRNA et RNU : (a) phase de validation, (b) phase de test.

#### IV.4.4 Analyse de sensibilité

Les résultats des contributions sont présentés dans la figure IV.7. Les variables les plus pertinentes, susceptibles d'influencer la prédiction des irradiances solaires, sont les suivantes : le numéro de Jours, direction du vent, mois, vitesse du vent, humidité relative, temps, température et pression atmosphérique. Toutes les variables d'entrée pertinentes citées dans cette contribution ont une contribution significative (IR > 9,2 %). L'utilisation de la méthode des poids dans l'analyse de sensibilité a efficacement révélé l'importance réelle de quasiment toutes les variables employées dans la prédiction de l'irradiation solaire globale.

## Chapitre IV : Prédiction de l'irradiation solaire globale horaire : comparaison des réseaux neuronaux / agrégation bootstrap



**Figure IV.7.** Illustration d'importance relative (IR) de chaque variable d'entrée par rapport à la sortie.

**Tableau IV.5.** Poids et biais de réseau neurone unique.

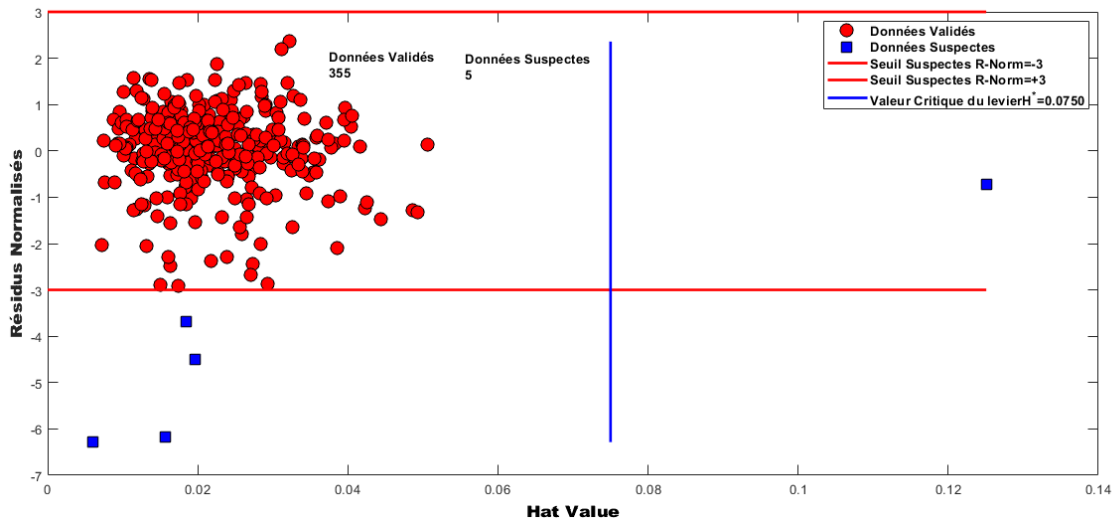
Connexions Couche d'entrée-Couche cachée								Connexions Couche cachée- Couche de sortie		
Poids								Biais	Poids	Biais
$w_{j1}^I$	$w_{j2}^I$	$w_{j3}^I$	$w_{j4}^I$	$w_{j5}^I$	$w_{j6}^I$	$w_{j7}^I$	$w_{j8}^I$	$b_j^H$	$w_{1j}^H$	$w_{1j}^H$
1.12146	-3.68556	-1.43416	2.13646	-1.47451	3.69746	-1.09522	0.02214	3.85698	-0.13720	0.43463
-0.70669	2.52181	-0.58307	-0.69459	0.52579	1.13093	-1.22758	0.13655	3.74836	-0.18745	
-2.79425	1.11464	0.76387	0.68959	0.27089	0.67938	-0.81162	-1.03988	3.11331	0.67448	
2.37071	-1.13932	0.37194	1.54417	0.03296	-2.54314	1.08934	0.40881	-1.83306	-0.15950	
1.69287	-2.90160	1.34819	-0.10445	-0.01886	-0.58755	1.49687	0.63111	-1.19867	0.12588	
-3.54506	-0.70765	-1.82749	-0.10193	-2.91451	0.12617	2.01129	1.52159	-0.86962	-0.08947	
-3.55923	1.24094	-0.45321	-2.90491	1.72737	0.01214	-1.78598	-0.38314	-0.01216	-0.23285	
-0.35685	3.39629	-3.94882	-0.71506	0.19187	0.01820	0.01693	0.70648	-1.82671	-0.81040	
-0.36812	3.27200	0.73179	-1.48871	-0.74427	0.85539	1.60865	-0.07912	-0.62301	-0.31853	
-0.83518	-0.10009	4.54292	1.51620	3.62047	-1.09928	0.42276	-1.83286	3.01769	-1.13343	
0.72500	0.44529	0.03770	0.15098	-0.39882	0.25842	-2.72557	-1.90079	-2.02250	0.98290	
-3.59354	1.96885	-0.14979	-0.27070	-3.94804	-1.61344	0.26132	0.72472	0.46348	1.08027	
-4.89343	5.51834	2.68023	-0.01722	2.33225	1.99865	0.70303	-1.02714	3.07894	-0.10903	
1.21481	-0.45819	0.06557	-0.21786	-0.00463	-0.44367	0.23290	1.34789	-0.64041	-0.95072	
0.75233	-1.14046	0.58189	-0.69028	-0.43128	0.19044	-0.23851	-0.26593	-1.03500	-0.28358	
-0.13428	1.26982	-0.78487	0.91526	-0.13848	-0.13890	1.17598	-0.79683	0.13439	-0.70895	
-0.36612	-0.74934	-0.15787	0.88482	0.43869	-0.24080	-2.90436	1.63886	0.54598	-1.72456	
-0.27835	-1.30158	-0.08963	0.93041	-1.03301	0.21593	0.24417	-2.46460	-2.61297	0.22546	
0.11706	-3.42139	-0.93354	1.49368	-0.46313	1.09742	1.11068	2.45742	-1.09886	0.17366	
1.08680	1.24363	-0.28412	0.23213	0.07652	0.50518	1.07963	-2.89952	-1.74578	0.10990	
0.11019	1.41482	-3.46622	-0.16541	-6.51842	-0.62077	0.83799	-2.38340	-0.78861	-1.57531	
0.58345	3.29992	1.47331	1.70738	0.17511	4.21169	0.91199	2.77241	2.46215	0.72611	
-0.26814	1.25320	-1.15928	-0.97202	0.46169	0.11031	-0.48862	2.03168	-0.13199	-0.07065	

#### IV.4.5 Diagnostic des points aberrants et domaine d'applicabilité du modèle neuronal optimisé

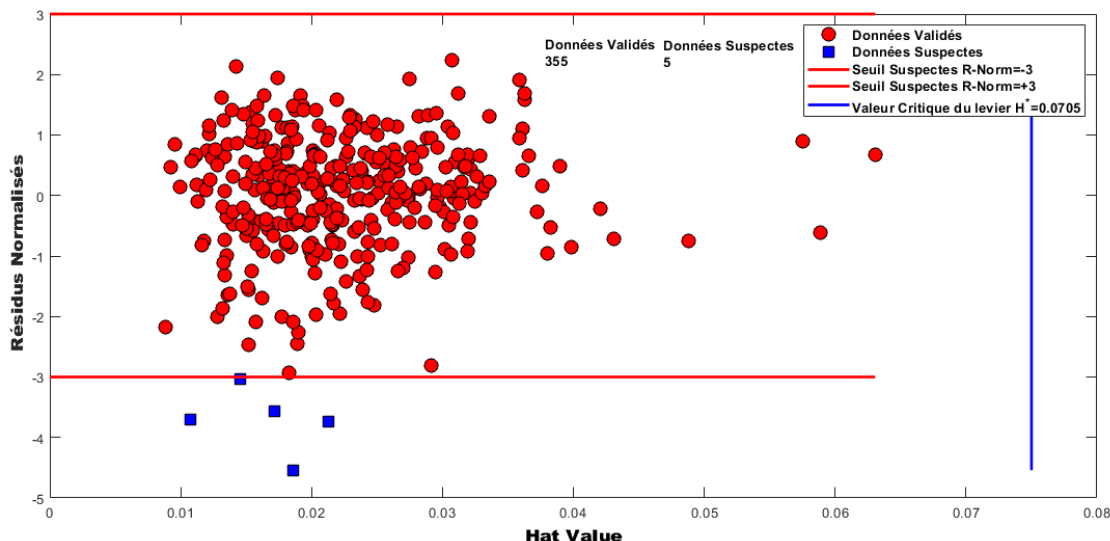
La figure IV.8 montre le diagramme de William pour le domaine d'applicabilité pour le modèle neuronal optimal "RNU" (Phase de validation et phase de test). Dans ce diagramme l'existence de 555/560 (98,61%) de points de données valides, 5/560 (1,39%) de points de données suspects. La valeur critique de levier

$$H^* = \frac{3(n + 1)}{m} = \frac{3(8 + 1)}{360} = 0.0750$$

Cela signifie que le développement du modèle RNU et ses prédictions sont réalisés dans le domaine d'applicabilité, ce qui aboutit à un modèle neuronal statistiquement valide et optimal. Par conséquent, nous pouvons affirmer qu'il existe des points de grande influence pour les deux phases



(a)



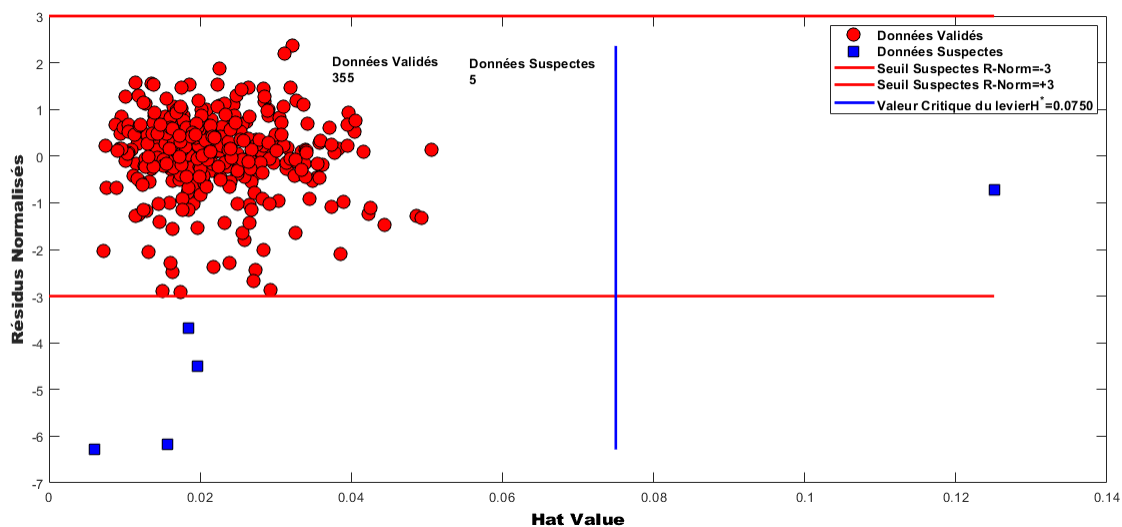
(b)

**Figure IV.8.** Diagnostic de données potentiellement suspectes et délimitation du domaine d'applicabilité du modèle neuronal "RNU" pour : (a) phase de validation, (b) phase de test.

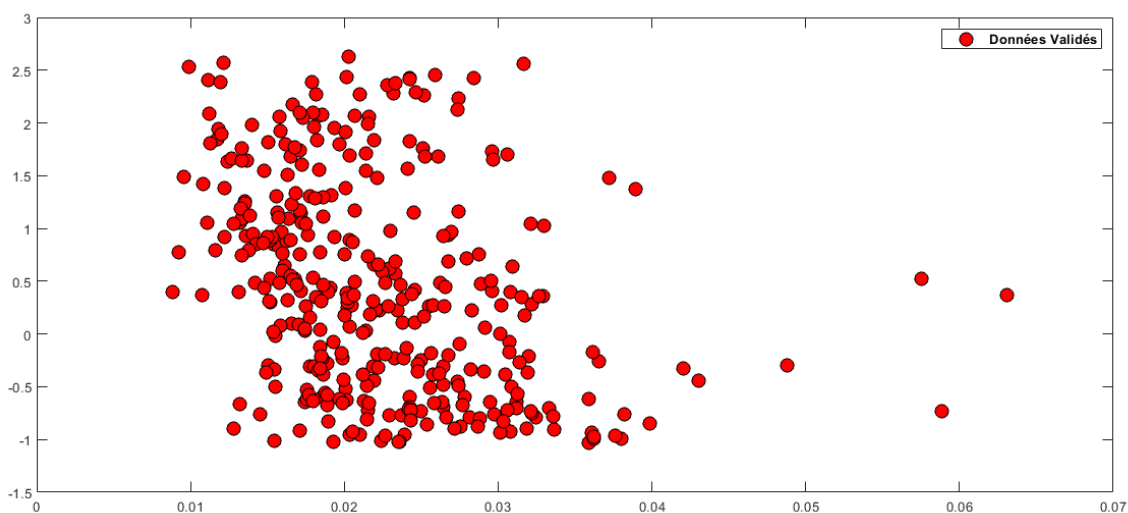
La figure IV.9 montre le diagramme de William pour le domaine d'applicabilité pour le modèle bootstrap réseaux de neurones agrégés "BRNA" (Phase de validation et phase de test). Dans ce diagramme l'existence de 555/560 (98,61%) de points de données valides, 5/560 (1,39%) de points de données suspectes. La valeur critique de levier

$$H^* = \frac{3(n+1)}{m} = \frac{3(8+1)}{360} = 0.0750$$

Pour la de validation et l'existence de 560/560 (100%) de points de données valides pour la phase de test. Cela indique que le développement du modèle BRNA et ses prédictions sont effectués dans le domaine d'applicabilité Ce qui conduit à un modèle bootstrap réseaux de neurones agrégés statistiquement valide. Ainsi on peut affirmer qu'il existe des points "Good High Leverage" pour les deux phases.



(a)



(b)

Figure IV.9. Diagnostic de données potentiellement suspectes et délimitation du domaine d'applicabilité du modèle neuronal "BRNA" pour : (a) phase de validation, (b) phase de test.

#### IV.4.6 Comparaison avec d'autres modèles

Pour évaluer l'importance des résultats obtenus, ils ont été comparés à des études similaires réalisées par d'autres chercheurs, en particulier aux modèles ayant utilisé les mêmes entrées que nous. Il a été constaté que, dans tous les modèles, l'objectif était de prédire l'irradiation solaire. Ces résultats confirment la robustesse et la précision du modèle BRNA pour l'irradiation solaire globale. Les résultats

obtenus à partir de ces modèles mentionnés, ainsi que les résultats de cette étude, sont présentés dans le tableau IV.6.

**Tableau VI.6.** Comparaison des résultats de notre modèle avec différents modèles similaires.

Références	Sites	Modèle	Indicateur d'évaluation
Dahmani et al. [103]	Bouzaréah (Algeria)	Bootstrap réseaux de neurones agrégés (BRNA)	R = 0.9680
Guermoui et al. [125]	Ghardaïa (Algeria)	Machine à vecteur de support (SVM)	R = 93.06
Linares-Rodríguez et al. [92]	Spain	Réseaux de neurones artificiels Perceptron à Multicouches (RNA-PMC)	R = 0.9400
Fadare. [126]	Nigeria	Perceptron à Multicouches feedforward, backpropagation (PMC-FBP)	R = 0.9560
Piri et al. [89]	Iran	Machine à vecteur de support pour régression (MVSR) et méthodes empiriques	R =0,9330

## IV.5 Conclusion

L'objectif de cette étude était de créer un modèle neuronal (RNAB) capable de prédire l'irradiation solaire horaire in Bouzaréah - (Algérie) basée sur les paramètres météorologiques.

Une comparaison des modèles RNAB et RNU a révélé que le modèle RNAB a donné les meilleurs résultats : La phase de validation avait une erreur quadratique moyenne de (REQM) 72,8758 Wh/m<sup>2</sup> pour RNAB et de 75,3748 Wh/m<sup>2</sup> pour RNU, alors que la phase de test présentait des erreurs quadratiques moyennes (REQM) de 62,4856 Wh/m<sup>2</sup> pour le RNAB et de 68,3968 Wh/m<sup>2</sup> pour le RNU. Le modèle RNAB a une plus grande précision et peut décrire la prédiction du l'irradiation solaire global horaire plus précise que le modèle RNU. Il confirme qu'un tel modèle de réseaux neuronaux agrégés bootstrap (RNAB) peut donner une prédiction plus précise et plus robuste qu'un modèle de réseaux neuronaux unique (RNU). Les prédictions robustes du modèle de réseaux neuronaux agrégés bootstrap ont été montrées par cette comparaison lorsqu'elles sont appliquées à des bases de données invisibles.



Pour conclure, notre modèle RNAB est adapté pour prédire l'irradiation solaire pour les lieux qui n'ont pas d'équipement de mesure (solarimètres/pyranomètres) et les systèmes concernés, car l'ensemble de données disponibles est petit, dans le cas d'un groupe de valeurs aberrantes qui doivent être exclues, données manquantes, ainsi que pour installer des systèmes d'énergie solaire et évaluer les conditions thermiques dans les études de construction en Algérie.

# Conclusion générale

## Conclusion générale

---

La prédiction de l'irradiation solaire joue un rôle fondamental dans les systèmes d'énergie solaire, l'irradiation solaire est la ressource énergétique initiale qui alimente tous les types de systèmes d'énergie solaire (Systèmes photovoltaïques (panneaux solaires), Systèmes solaires thermiques, Concentration solaire, Systèmes solaires hybrides...etc.). Son niveau et sa disponibilité varient en fonction de la géolocalisation, de la météo et d'autres facteurs environnementaux. Par conséquent, la conception, la performance et l'efficacité des systèmes d'énergie solaire dépendent en grande partie de la quantité et de la qualité de l'irradiation solaire reçue sur la surface de la terre.

Les techniques d'intelligence artificielle (IA) possèdent une caractéristique exceptionnelle dont l'utilité se manifeste lors de la représentation des procédés non linéaires : elles agissent comme des approximations universelles concises. Les techniques d'intelligence artificielle (IA) comme (les réseaux de neurones artificiels (RNA) et Machine à vecteur support (MVS)) ont finalement été appliqués pour la prédiction de l'irradiation solaire et présentent un intérêt accru dans la modélisation des systèmes d'énergie solaire.

En fait, Le but de cette thèse consiste à créer des modèles optimaux par les techniques d'intelligence artificielle (IA) qui pourra être utilisé pour prédire l'irradiation solaire globale horaire reçue sur un plan horizontal. Cette démarche de modélisation se déploie au travers du développement d'un programme sur les logiciels STATISTICA et MATLAB 2020b.

La première contribution de cette étude réside dans l'utilisation de modèles conventionnels basés sur l'apprentissage automatique traditionnel, notamment les réseaux de neurones artificiels de type Perceptron à Multicouches (RNA-PMC) et les machines à vecteur support avec la fonction noyau radiale à base (MVS-RBF). Ces modèles ont été employés pour prédire l'irradiation solaire horaire en se basant sur des paramètres d'entrée mesurables dans deux stations radiométriques situées respectivement à Bouzaréah et Tamanrasset, dans le nord et le sud de l'Algérie. L'indice d'évaluation comme le coefficient de corrélation (R) l'erreur moyenne en pourcentage (MEP) et la racine de l'erreur quadratique moyenne (REQM) ont été déterminés et comparés avec d'autres modèles simulés de l'étude précédentes. Il a été constaté que le modèle (MVS-RBF) a montré une forte corrélation dans la stratégie utilisée, avec un coefficient de corrélation R de 0,9876, une erreur moyenne prédictive (MEP) de 8,1586 % et une racine de l'erreur quadratique moyenne (REQM) de 38,7045 Wh/m<sup>2</sup>, respectivement pour le site de Bouzaréah.

Dans la deuxième contribution, on a développé un modèle bootstrap réseaux de neurones agrégés (BRNA) pour prédire l'irradiation solaire globale dans la cité de Bouzaréah–Algérie. On note qu'un

## Conclusion générale

---

coefficient de corrélation de 96.80 %. L'erreur moyenne prédictive "EMP" dans le modèle développé BRNA atteint 12.96 %. En plus, le modèle donne une racine de l'erreur quadratique moyenne (REQM) de 62.4856 Wh/m<sup>2</sup>.

Ces résultats montrent que les modèles optimisés (SVM-RBF et BRNA) sont résilients, avec une forte capacité prédictive soutenue par une concordance satisfaisante entre les valeurs mesurées et prédites de l'irradiation solaire globale horaire. Ce travail est porté sur les données météorologiques horaires (Mois, Numéro de jour, Temps, Température, Relative humidité, Pression atmosphérique, Vitesse de vent, Direction de vent) pour les villes de Bouzaréah et Tamanrasset, Il serait envisageable d'étendre cette analyse aux zones climatiques présentant des conditions météorologiques similaires.

De plus, la comparaison des résultats entre le modèle du machine à vecteur support avec la fonction de noyau radiale à base (MVS-RBF) et le modèle les réseaux de neurones artificiels type de perceptron à multicouches (RNA-PMC) a montré la supériorité du modèle (MVS-RBF) et une autre comparaison des résultats entre notre modèle développée (BRNA) et le modèle du réseaux de neurones unique (RNU) a montré la supériorité du modèle (BRNA). L'analyse de sensibilité a été réalisée comme une technique pour déterminer la contribution des paramètres d'entrée à l'irradiation solaire globale en appliquant la méthode des poids de connexion. Il est à noter que les paramètres les plus importants qui peuvent influencer sur l'irradiation solaire globale pour l'obtention de résultats de haute précision sont : La pression (16%), température, jour, l'humidité relative avec une importance relative (IR) de 13% pour la première contribution, et jour de (16%), direction de vent, moins et vitesse de vent avec RI = 14% pour la deuxième contribution.

Comme perspectives à notre étude nous envisageons :

- D'appliquer les techniques d'intelligence artificielle (IA) d'autres villes sous conditions climatiques algériennes simulées pour prédire l'irradiation solaire globale sur des surfaces (horizontales-inclinés).
- D'utiliser d'autres types améliorés comme bootstrap réseaux de neurones agrégés (BRNA) pour la prédiction de l'irradiation solaire globale, ou des méthodes hybrides par exemple entre la logique floue avec les réseaux de neurones.
- D'utiliser une carte de l'irradiation solaire disponible sur l'Algérie pour les applications de l'énergie solaire, car elle peut illustrer les endroits optimaux pour monter les installations des systèmes d'énergie solaire.

## **Conclusion générale**

---

Nous souhaitons que cette étude puisse stimuler la recherche et les investigations des gisements solaires, et le potentiel d'autres sources d'énergie propre : l'énergie de biomasse, éolienne, et la géothermie.

# Références bibliographiques

## Références bibliographiques

---

- [1] B. Amiri, “Estimation temporelle du rayonnement solaire au nord et au Sahara Algérien à partir de données de mesures sélectionnées,” 2021.
- [2] S. N. Mughal, Y. R. Sood, and R. K. Jarial, “A novel hybrid model for predicting hourly global solar radiations on the tilted surface,” *Int. J. Environ. Sci. Technol.*, 2022, doi: 10.1007/s13762-021-03903-0.
- [3] F. Besharat, A. A. Dehghan, and A. R. Faghieh, “Empirical models for estimating global solar radiation: A review and case study,” *Renew. Sustain. Energy Rev.*, vol. 21, pp. 798–821, 2013, doi: 10.1016/j.rser.2012.12.043.
- [4] A. K. and Katiyar and C. K. Pandey, “Simple correlation for estimating the global solar radiation on horizontal surfaces in India,” *Energy*, vol. 35, no. 12, pp. 5043–5048, 2010.
- [5] N. Bailek *et al.*, “A new empirical model for forecasting the diffuse solar radiation over Sahara in the Algerian Big South,” *Renew. Energy*, vol. 117, pp. 530–537, 2018, doi: 10.1016/j.renene.2017.10.081.
- [6] B. Amiri, A. M. Gómez-Orellana, P. A. Gutiérrez, R. Dizène, C. Hervás-Martínez, and K. Dahmani, “A novel approach for global solar irradiation forecasting on tilted plane using Hybrid Evolutionary Neural Networks,” *J. Clean. Prod.*, vol. 287, Mar. 2021, doi: 10.1016/j.jclepro.2020.125577.
- [7] M. Guermoui, K. Bouchouicha, S. Benkaciali, K. Gairaa, and N. Bailek, “New soft computing model for multi-hours forecasting of global solar radiation,” *Eur. Phys. J. Plus*, vol. 137, no. 1, Jan. 2022, doi: 10.1140/epjp/s13360-021-02263-5.
- [8] R. Dubayah and P. M. Rich, “Topographic solar radiation models for GIS,” *Int. J. Geogr. Inf. Syst.*, vol. 9, no. 4, pp. 405–419, 1995.
- [9] P. Aguayo, “Solar energy potential analysis at building scale using LiDAR and satellite data.” University of Waterloo, 2013.
- [10] S. Gadsden, M. Rylatt, K. Lomas, and D. Robinson, “Predicting the urban solar fraction: a methodology for energy advisers and planners based on GIS,” *Energy Build.*, vol. 35, no. 1, pp. 37–48, 2003.
- [11] M. Šúri and J. Hofierka, “A new GIS-based solar radiation model and its application to photovoltaic assessments,” *Trans. GIS*, vol. 8, no. 2, pp. 175–190, 2004.
- [12] J. A. Duffie and W. A. Beckman, *Solar engineering of thermal processes*. Wiley New York, 1980.
- [13] H. E. Newell, *Beyond the atmosphere: Early years of space science*. Courier Corporation, 2011.
- [14] R. W. Mueller, “Solar irradiance, global distribution,” in *Solar energy*, Springer, 2013, pp. 553–583.
- [15] K. Dahmani, “Contribution au développement de méthodes numériques utilisant les réseaux de neurones artificiels pour la caractérisation du gisement solaire en Algérie.” Université des sciences et de la technologie Houari-Boumediène, 2015.
- [16] S. C. Bhatia, *Advanced renewable energy systems, (Part 1 and 2)*. CRC Press, 2014.
- [17] CHAIBI MOHAMED, “Gisements solaires et éoliens au Maroc : estimation et évaluation par intelligence artificielle et systèmes d’information géographique,” PES. Faculté des Sciences- Meknès, 2022.
- [18] B. E. Psiloglou and H. D. Kambezidis, “Performance of the meteorological radiation model during the solar eclipse of 29 March 2006,” *Atmos. Chem. Phys.*, vol. 7, no. 23, pp. 6047–6059, 2007.
- [19] A. Rezrazi, “Contribution à la modélisation de l’irradiation solaire,” Université de Médea, 2016.
- [20] V. Badescu, “Modeling solar radiation at the earth’s surface,” *SpringerVerlag, Berlin/heidelb.*, 2008.

## Références bibliographiques

---

- [21] A. Oudrane, “Contribution to the Modeling and Development of Solar Heating Systems for Individual Use.” Ecole Nationale Polytechnique d’Oran Maurice Audin, 2018.
- [22] G. N. Tiwari and A. Tiwari, *Handbook of solar energy*, vol. 498. Springer, 2016.
- [23] S. CHABA MOUNA and S. HANINI, “Usage des réseaux de neurones pour le calcul des paramètres énergétiques solaires.” 2018.
- [24] R. Foster, M. Ghassemi, and A. Cota, *Solar energy: renewable energy and the environment*. CRC press, 2009.
- [25] M. Iqbal, *An introduction to solar radiation*. Elsevier, 2012.
- [26] S. V Szokolay, “Solar Geometry—Passive and Low Energy, Architecture International in Association with Department of Architecture,” *Univ. Queensland, Brisbane, Aust.*, 1996.
- [27] C. A. Gueymard, D. Myers, and K. Emery, “Proposed reference irradiance spectra for solar energy systems testing,” *Sol. energy*, vol. 73, no. 6, pp. 443–467, 2002.
- [28] M. Paulescu, E. Paulescu, P. Gravila, and V. Badescu, *Weather modeling and forecasting of PV systems operation*, vol. 358. Springer, 2013.
- [29] S. Wilbert *et al.*, “Measuring solar radiation and relevant atmospheric parameters,” 2017.
- [30] C. A. Gueymard, “Parameterized transmittance model for direct beam and circumsolar spectral irradiance,” *Sol. Energy*, vol. 71, no. 5, pp. 325–346, 2001.
- [31] P. Ineichen, “Comparison of eight clear sky broadband models against 16 independent data banks,” *Sol. Energy*, vol. 80, no. 4, pp. 468–478, 2006.
- [32] N. Abdelhafidi, N. E. I. Bachari, and Z. Abdelhafidi, “Estimation of solar radiation using stepwise multiple linear regression with principal component analysis in Algeria,” *Meteorol. Atmos. Phys.*, vol. 133, pp. 205–216, 2021.
- [33] G. S. ATLAS, “An innovation of the world Bank Group. 2016. www.” 2017.
- [34] Y. Li *et al.*, “Climate model shows large-scale wind and solar farms in the Sahara increase rain and vegetation,” *Science (80-. )*, vol. 361, no. 6406, pp. 1019–1022, 2018.
- [35] B. Saïd, “Prédiction du DNI en Algérie : cas des modèles hybrides,” Ecole Nationale Polytechnique, 2019. [Online]. Available: [www.enp.edu.dz](http://www.enp.edu.dz)
- [36] M. R. Yaiche, A. Bouhanik, S. M. A. Bekkouche, A. Malek, and T. Benouaz, “Revised solar maps of Algeria based on sunshine duration,” *Energy Convers. Manag.*, vol. 82, pp. 114–123, 2014.
- [37] “Ministère de l’Énergie | Algérie,” 2011. <https://www.energy.gov.dz/>
- [38] K. FARKOUS, “Caractérisation et évaluation du rayonnement solaire pour le site de Ghardaïa.” Université de Djelfa-Ziane Achour.
- [39] J.-P. Haton and M.-C. Haton, “L’intelligence artificielle,” (*No Title*), 1989.
- [40] D. I. Poole, R. G. Goebel, and A. K. Mackworth, *Computational intelligence*, vol. 1. Oxford University Press Oxford, 1998.
- [41] L. R. Medsker, “Microcomputer applications of hybrid intelligent systems,” *J. Netw. Comput. Appl.*, vol. 19, no. 2, pp. 213–234, 1996.
- [42] H. BOUZEMMAL, “MODELISATION DES PARAMETRES METEOROLOGIQUES. ETUDE



- COMPARATIVE.” 2023.
- [43] C. Si Moussa, “Modélisation des équilibres de phases à pressions élevées par les réseaux de neurones artificiels.” 2010.
- [44] Ammi, “Modelisation des phenomenes de transport lors d’osmose inverse/nanofiltration,” Université Yahia Fares de Médéa.
- [45] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: theory and applications,” *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
- [46] S. Belaid and A. Mellit, “Prediction of daily and mean monthly global solar radiation using support vector machine in an arid climate,” *Energy Convers. Manag.*, vol. 118, pp. 105–118, 2016.
- [47] L. Bottou and O. Bousquet, “The tradeoffs of large scale learning,” *Adv. Neural Inf. Process. Syst.*, vol. 20, 2007.
- [48] J. García-Alba, J. F. Bárcena, C. Ugarteburu, and A. García, “Artificial neural networks as emulators of process-based models to analyse bathing water quality in estuaries,” *Water Res.*, vol. 150, pp. 283–295, 2019, doi: 10.1016/j.watres.2018.11.063.
- [49] C. N. OBIORA, “Deep Learning Approach To Forecasting Hourly Solar Irradiance,” *Вестник Росздравнадзора*, vol. 6, no. August, pp. 5–9, 2017.
- [50] K. Ho, N. KO, I. T. U. IH, and B. ÉL, “Tim.«,” *Random Decis. For. AT&T Bell Lab.*, pp. 278–282, 1995.
- [51] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
- [52] J. Fan *et al.*, “Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China,” *Energy Convers. Manag.*, vol. 164, pp. 102–111, 2018.
- [53] A. Khamis and S. Abdullah, “Forecasting wheat price using backpropagation and NARX neural network,” *Int. J. Eng. Sci.*, vol. 3, no. 11, pp. 19–26, 2014.
- [54] E. Pisoni, M. Farina, C. Carnevale, and L. Piroddi, “Forecasting peak air pollution levels using NARX models,” *Eng. Appl. Artif. Intell.*, vol. 22, no. 4–5, pp. 593–602, 2009.
- [55] S. Mohanty, P. K. Patra, and S. S. Sahoo, “Prediction of global solar radiation using nonlinear autoregressive network with exogenous inputs (narx),” in *2015 39th National Systems Conference (NSC)*, 2015, pp. 1–6.
- [56] G. U. Yule, “On a method of investigating periodicities in disturbed series with special reference to Wolfer’s sunspot numbers,” *Stat. Pap. Georg. Udry Yule*, pp. 389–420, 1971.
- [57] G. E. P. Box, G. M. Jenkins, and J. F. MacGregor, “Some recent advances in forecasting and control,” *J. R. Stat. Soc. Ser. C (Applied Stat.)*, vol. 23, no. 2, pp. 158–179, 1974.
- [58] J. Wu, “Statistical and data mining approach for the prediction of solar radiation.” 2013.
- [59] W. ZELLNER, *An introduction to Bayesian inference in econometrics*. 1996.
- [60] E. Parzen, “ARARMA models for time series analysis and forecasting,” *J. Forecast.*, vol. 1, no. 1, pp. 67–82, 1982.
- [61] J. G. de Gooijer and A. Klein, “On the cumulated multi-step-ahead predictions of vector autoregressive moving average processes,” *Int. J. Forecast.*, vol. 7, no. 4, pp. 501–513, 1992.

## Références bibliographiques

---

- [62] M. J. Artis and W. Zhang, “BVAR forecasts for the G-7,” *Int. J. Forecast.*, vol. 6, no. 3, pp. 349–362, 1990.
- [63] Z. Wang and D. A. Bessler, “Forecasting performance of multivariate time series models with full and reduced rank: An empirical examination,” *Int. J. Forecast.*, vol. 20, no. 4, pp. 683–695, 2004.
- [64] S. Hokoï, M. Matsumoto, and M. Kagawa, “Stochastic models of solar radiation and outdoor temperature,” *ASHRAE Trans. (American Soc. Heating, Refrig. Air-Conditioning Eng. States)*, vol. 96, no. CONF-9006117-, 1990.
- [65] B. H. Chowdhury and S. Rahman, “Forecasting sub-hourly solar irradiance for prediction of photovoltaic output,” in *19th ieee photovoltaic specialists conference*, 1987, pp. 171–176.
- [66] L. L. Mora-Lopez and M. Sidrach-de-Cardona, “Multiplicative ARMA models to generate hourly series of global irradiation,” *Sol. Energy*, vol. 63, no. 5, pp. 283–291, 1998.
- [67] Y. Jiang, “Computation of monthly mean daily global solar radiation in China using artificial neural networks and comparison with other empirical models,” *Energy*, vol. 34, no. 9, pp. 1276–1283, 2009.
- [68] J. G. Fortin, F. Anctil, L.-É. Parent, and M. A. Bolinder, “Comparison of empirical daily surface incoming solar radiation models,” *Agric. For. Meteorol.*, vol. 148, no. 8–9, pp. 1332–1340, 2008.
- [69] K. Dahmani, R. Dizene, G. Notton, C. Paoli, C. Voyant, and M. L. Nivet, “Estimation of 5-min time-step data of tilted solar global irradiation using ANN (Artificial Neural Network) model,” *Energy*, vol. 70, pp. 374–381, 2014, doi: 10.1016/j.energy.2014.04.011.
- [70] T. Khatib, A. Mohamed, K. Sopian, and M. Mahmoud, “Assessment of artificial neural networks for hourly solar radiation prediction,” *Int. J. Photoenergy*, vol. 2012, 2012, doi: 10.1155/2012/946890.
- [71] S. X. Chen, H. B. Gooi, and M. Q. Wang, “Solar radiation forecast based on fuzzy logic and neural networks,” *Renew. energy*, vol. 60, pp. 195–201, 2013.
- [72] N. D. Kaushika, R. K. Tomar, and S. C. Kaushik, “Artificial neural network model based on interrelationship of direct, diffuse and global solar radiations,” *Sol. Energy*, vol. 103, pp. 327–342, 2014.
- [73] M. Laidi, S. Hanini, A. Rezrazi, M. R. Yaiche, A. A. El Hadj, and F. Chellali, “Supervised artificial neural network-based method for conversion of solar radiation data (case study: Algeria),” *Theor. Appl. Climatol.*, vol. 128, no. 1–2, pp. 439–451, 2017, doi: 10.1007/s00704-015-1720-7.
- [74] S. Mehdizadeh, J. Behmanesh, and K. Khalili, “Comparison of artificial intelligence methods and empirical equations to estimate daily solar radiation,” *J. Atmos. Solar-Terrestrial Phys.*, vol. 146, pp. 215–227, 2016.
- [75] A. Marzo *et al.*, “Daily global solar radiation estimation in desert areas using daily extreme temperatures and extraterrestrial radiation,” *Renew. Energy*, vol. 113, pp. 303–311, 2017.
- [76] V. H. Quej, J. Almorox, J. A. Arnaldo, and L. Saito, “ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment,” *J. Atmos. Solar-Terrestrial Phys.*, vol. 155, pp. 62–70, 2017.
- [77] A. T. Mohammad, I. Mahmood, and S. Mahjoob, “Three structures of a multilayer artificial neural network for predicting the solar radiation of Baghdad City-Iraq,” *Int. J. Appl. Eng. Res.*, vol. 13, no. 6, pp. 3759–3766, 2018.
- [78] L. Benali, G. Notton, A. Fouilloy, C. Voyant, and R. Dizene, “Solar radiation forecasting using artificial neural network and random forest methods: Application to normal beam, horizontal diffuse and global components,” *Renew. energy*, vol. 132, pp. 871–884, 2019.

## Références bibliographiques

---

- [79] H. Sharadga, S. Hajimirza, and R. S. Balog, “Time series forecasting of solar power generation for large-scale photovoltaic plants,” *Renew. Energy*, vol. 150, pp. 797–807, 2020.
- [80] R. Kumar, R. K. Aggarwal, and J. D. Sharma, “Comparison of regression and artificial neural network models for estimation of global solar radiations,” *Renew. Sustain. Energy Rev.*, vol. 52, pp. 1294–1299, 2015.
- [81] M. T. Ismail, N. Z. A. Shah, and S. A. A. Karim, “Modeling solar radiation in peninsular Malaysia using ARIMA model,” *Clean Energy Oppor. Trop. Ctries.*, pp. 53–71, 2021.
- [82] Z. Bounoua, L. O. Chahidi, and A. Mechaqrane, “Estimation of daily global solar radiation using empirical and machine-learning methods: A case study of five Moroccan locations,” *Sustain. Mater. Technol.*, vol. 28, p. e00261, 2021.
- [83] M. Sridharan, “Generalized regression neural network model based estimation of global solar energy using meteorological parameters,” *Ann. Data Sci.*, vol. 10, no. 4, pp. 1107–1125, 2023.
- [84] A. Mellit, M. Menghanem, and M. Bendekhis, “Artificial neural network model for prediction solar radiation data: application for sizing stand-alone photovoltaic power system,” in *IEEE Power Engineering Society General Meeting, 2005*, 2005, pp. 40–44.
- [85] R. Yacef, A. Mellit, S. Belaid, and Z. Şen, “New combined models for estimating daily global solar radiation from measured air temperature in semi-arid climates: application in Ghardaïa, Algeria,” *Energy Convers. Manag.*, vol. 79, pp. 606–615, 2014.
- [86] S. C. Nwokolo and J. C. Ogbulezie, “A quantitative review and classification of empirical models for predicting global solar radiation in West Africa,” *Beni-Suef Univ. J. basic Appl. Sci.*, vol. 7, no. 4, pp. 367–396, 2018.
- [87] T. R. Ayodele and A. S. O. Ogunjuyigbe, “Performance assessment of empirical models for prediction of daily and monthly average global solar radiation: the case study of Ibadan, Nigeria,” *Int. J. Ambient Energy*, vol. 38, no. 8, pp. 803–813, 2017.
- [88] K. Benmouiza and A. Cheknane, “Forecasting hourly global solar radiation using hybrid k -means and nonlinear autoregressive neural network models,” *Energy Convers. Manag.*, vol. 75, pp. 561–569, 2013, doi: 10.1016/j.enconman.2013.07.003.
- [89] J. Piri, S. Shamshirband, D. Petković, C. W. Tong, and M. H. ur Rehman, “Prediction of the solar radiation on the Earth using support vector regression technique,” *Infrared Phys. & Technol.*, vol. 68, pp. 179–185, 2015, doi: 10.1016/j.infrared.2014.12.006.
- [90] H. Suyono, H. Santoso, R. N. Hasanah, U. Wibawa, and I. Musirin, “Prediction of solar radiation intensity using extreme learning machine,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 12, no. 2, pp. 691–698, 2018.
- [91] M. Benghanem and A. Mellit, “Radial Basis Function Network-based prediction of global solar radiation data: Application for sizing of a stand-alone photovoltaic system at Al-Madinah, Saudi Arabia,” *Energy*, vol. 35, no. 9, pp. 3751–3762, 2010, doi: 10.1016/j.energy.2010.05.024.
- [92] A. Linares-Rodriguez, J. A. Ruiz-Arias, D. Pozo-Vazquez, and J. Tovar-Pescador, “An artificial neural network ensemble model for estimating global solar radiation from Meteosat satellite images,” *Energy*, vol. 61, pp. 636–645, 2013.
- [93] A. Koca, H. F. Oztop, Y. Varol, and G. O. Koca, “Estimation of solar radiation using artificial neural networks with different input parameters for Mediterranean region of Anatolia in Turkey,” *Expert Syst. Appl.*, vol. 38, no. 7, pp. 8756–8762, 2011.

## Références bibliographiques

---

- [94] C. Voyant, C. Paoli, M. Muselli, and M.-L. Nivet, “Multi-horizon solar radiation forecasting for Mediterranean locations using time series models,” *Renew. Sustain. Energy Rev.*, vol. 28, pp. 44–52, 2013.
- [95] C. M. Siham, H. Salah, L. Maamar, and K. Latifa, “Artificial neural networks based prediction of hourly horizontal solar radiation data: Case study,” *Int. J. Appl. Decis. Sci.*, vol. 10, no. 2, pp. 156–174, 2017, doi: 10.1504/IJADS.2017.084312.
- [96] M. Castangia, A. Aliberti, L. Bottaccioli, E. Macii, and E. Patti, “A compound of feature selection techniques to improve solar radiation forecasting,” *Expert Syst. Appl.*, vol. 178, p. 114979, 2021.
- [97] L. Huang, J. Kang, M. Wan, L. Fang, C. Zhang, and Z. Zeng, “Solar radiation prediction using different machine learning algorithms and implications for extreme climate events,” *Front. Earth Sci.*, vol. 9, p. 596860, 2021.
- [98] M. Guermoui, R. Abdelaziz, K. Gairaa, L. Djemoui, and S. Benkaciali, “New temperature-based predicting model for global solar radiation using support vector regression,” *Int. J. Ambient Energy*, vol. 43, no. 1, pp. 1397–1407, 2022, doi: 10.1080/01430750.2019.1708792.
- [99] O. Bamisile, A. Oluwasanmi, C. Ejayi, N. Yimen, S. Obiora, and Q. Huang, “Comparison of machine learning and deep learning algorithms for hourly global/diffuse solar radiation predictions,” *Int. J. Energy Res.*, vol. 46, no. 8, pp. 10052–10073, 2022.
- [100] İ. Üstün, F. Üneş, İ. Mert, and C. Karakuş, “A comparative study of estimating solar radiation using machine learning approaches: DL, SMGRT, and ANFIS,” *Energy Sources, Part A Recover. Util. Environ. Eff.*, vol. 44, no. 4, pp. 10322–10345, 2022.
- [101] Y. El Mghouchi, “Best Combinations of Inputs for ANN-Based Solar Radiation Forecasting in Morocco,” *Technol. Econ. Smart Grids Sustain. Energy*, vol. 7, no. 1, Dec. 2022, doi: 10.1007/s40866-022-00152-z.
- [102] Y. Ammi, L. Khaouane, and S. Hanini, “Stacked neural networks for predicting the membranes performance by treating the pharmaceutical active compounds,” *Neural Comput. Appl.*, vol. 33, no. 19, pp. 12429–12444, 2021, doi: 10.1007/s00521-021-05876-0.
- [103] A. Dahmani, Y. Ammi, S. Hanini, M. Redha Yaiche, and H. Zentou, “Prediction of Hourly Global Solar Radiation: Comparison of Neural Networks/Bootstrap Aggregating,” *Kem. u Ind. Časopis kemičara i Kem. inženjera Hrvat.*, vol. 72, no. 3–4, pp. 201–213, 2023.
- [104] A. Takilalte, S. Harrouni, M. R. Yaiche, and L. Mora-López, “New approach to estimate 5-min global solar irradiation data on tilted planes from horizontal measurement,” *Renew. Energy*, vol. 145, pp. 2477–2488, 2020, doi: 10.1016/j.renene.2019.07.165.
- [105] M. Zaid, S. Khouildat, and A. R. Siagh, “Strategy of the renewable energy in Algeria, as an inevitable drift to diversification of the economy,” *Transport*, vol. 13, no. 13,889, pp. 3–9, 2017.
- [106] A. Rezrazi, S. Hanini, and M. Laidi, “An optimisation methodology of artificial neural network models for predicting solar radiation: a case study,” *Theor. Appl. Climatol.*, vol. 123, no. 3–4, pp. 769–783, 2016, doi: 10.1007/s00704-015-1398-x.
- [107] A. Dahmani *et al.*, “Assessing the Efficacy of Improved Learning in Hourly Global Irradiance Prediction,” *Comput. Mater. Contin.*, vol. 77, no. 2, 2023.
- [108] G. D. Garson, “Interpreting neural-network connection weights,” *AI Expert*, vol. 6, no. 4, pp. 46–51, 1991.

## Références bibliographiques

---

- [109] A. T. C. Goh, “Back-propagation neural networks for modeling complex systems,” *Artif. Intell. Eng.*, vol. 9, no. 3, pp. 143–151, 1995.
- [110] M. Gevrey, I. Dimopoulos, and S. Lek, “Review and comparison of methods to study the contribution of variables in artificial neural network models,” *Ecol. Modell.*, vol. 160, no. 3, pp. 249–264, 2003, doi: 10.1016/S0304-3800(02)00257-0.
- [111] Y. Ammi, L. Khaouane, and S. Hanini, “Prediction of the rejection of organic compounds (neutral and ionic) by nanofiltration and reverse osmosis membranes using neural networks,” *Korean J. Chem. Eng.*, vol. 32, no. 11, pp. 2300–2310, 2015, doi: 10.1007/s11814-015-0086-y.
- [112] M. Hosseinzadeh and A. Hemmati-Sarapardeh, “Toward a predictive model for estimating viscosity of ternary mixtures containing ionic liquids,” *J. Mol. Liq.*, vol. 200, pp. 340–348, 2014.
- [113] J.-L. Chen *et al.*, “Estimation of monthly-mean global solar radiation using MODIS atmospheric product over China,” *J. Atmos. Solar-Terrestrial Phys.*, vol. 110, pp. 63–80, 2014.
- [114] S. Bamehr and S. Sabetghadam, “Estimation of global solar radiation data based on satellite-derived atmospheric parameters over the urban area of Mashhad, Iran,” *Environ. Sci. Pollut. Res.*, vol. 28, no. 6, pp. 7167–7179, 2021.
- [115] İ. Uçkan and K. M. Khudhur, “Improving of global solar radiation forecast by comparing other meteorological parameter models with sunshine duration models,” *Environ. Sci. Pollut. Res.*, vol. 29, no. 25, pp. 37867–37881, 2022, doi: 10.1007/s11356-022-18781-3.
- [116] R. Marquez and C. F. M. Coimbra, “Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database,” *Sol. Energy*, vol. 85, no. 5, pp. 746–756, 2011.
- [117] P. A. C. Rocha and V. O. Santos, “Global horizontal and direct normal solar irradiance modeling by the machine learning methods XGBoost and deep neural networks with CNN-LSTM layers: a case study using the GOES-16 satellite imagery,” *Int. J. Energy Environ. Eng.*, pp. 1–16, 2022.
- [118] S. A. Kalogirou, E. Mathioulakis, and V. Belessiotis, “Artificial neural networks for the performance prediction of large solar systems,” *Renew. Energy*, vol. 63, pp. 90–97, 2014.
- [119] L. Khaouane, Y. Ammi, and S. Hanini, “Modeling the retention of organic compounds by nanofiltration and reverse osmosis membranes using bootstrap aggregated neural networks,” *Arab. J. Sci. Eng.*, vol. 42, no. 4, pp. 1443–1453, 2017.
- [120] J. Zhang, “Inferential estimation of polymer quality using bootstrap aggregated neural networks,” *Neural networks*, vol. 12, no. 6, pp. 927–938, 1999.
- [121] R. J. Tibshirani and B. Efron, “An introduction to the bootstrap,” *Monogr. Stat. Appl. Probab.*, vol. 57, pp. 1–436, 1993.
- [122] Y. Ammi, L. Khaouane, and S. Hanini, “A Model Based on Bootstrapped Neural Networks for Modeling the Removal of Organic Compounds by Nanofiltration and Reverse Osmosis Membranes,” *Arab. J. Sci. Eng.*, vol. 43, no. 11, pp. 6271–6284, 2018, doi: 10.1007/s13369-018-3484-8.
- [123] J. Kiseľák, Y. Lu, J. Švihra, P. Szépe, and M. Stehlik, “‘SPOCU’: scaled polynomial constant unit activation function,” *Neural Comput. Appl.*, vol. 33, no. 8, pp. 3385–3401, 2021.
- [124] Y. Ammi, L. Khaouane, and S. Hanini, “Stacked neural networks for predicting the membranes performance by treating the pharmaceutical active compounds,” *Neural Comput. Appl.*, vol. 33, no. 19, pp. 12429–12444, 2021, doi: 10.1007/s00521-021-05876-0.

## Références bibliographiques

---

- [125] M. Guermoui and A. Rabehi, “Soft computing for solar radiation potential assessment in Algeria,” *Int. J. Ambient Energy*, vol. 41, no. 13, pp. 1524–1533, 2020.
- [126] D. A. Fadare, “Modelling of solar energy potential in Nigeria using an artificial neural network model,” *Appl. Energy*, vol. 86, no. 9, pp. 1410–1422, 2009, doi: 10.1016/j.apenergy.2008.12.005.