

الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

جامعة أبي بكر بلقايد - تلمسان

Université Aboubakr Belkaïd – Tlemcen –

Faculté de TECHNOLOGIE



MEMOIRE

Présenté pour l'obtention du **diplôme** de **MASTER**

En : Electrotechnique

Spécialité : Commandes Electriques

Par : Ouahrani Mohammed El Amine & Riahi Mohammed Salim

Sujet

Prédiction de la production d'énergie renouvelable en utilisant
l'intelligence artificielle

Soutenu publiquement, le 02 / 06 / 2025 , devant le jury composé de :

M. Dris Younes	MCB	Université de Tlemcen	Président
M. Mechernene Abdelkader	MCA	Université de Tlemcen	Examinateur
Mme Bourri Sihem	MCA	Université de Tlemcen	Encadreur
M. Messaoud Hamouda	Pr	Université d'Adrar	Co-Encadreur

Année universitaire : 2024 /2025

Dedicaces 1

Je dédie ce modeste travail à ma mère et mon père, symboles d'affection, de patience et d'amour inconditionnel, qui ont sacrifié leur temps et leur énergie pour mon bonheur et ma réussite. Qu'ils trouvent ici l'expression de ma profonde gratitude et de mon respect éternel.

Je pense également avec une immense tendresse à tous les membres de ma Famille — mes oncles, mes tantes, mes cousins et cousines — pour leur présence constante, leurs encouragements et leur bienveillance tout au long de mes études.

À mes amis et collègues, merci pour votre soutien, vos partages, et les Moments précieux vécus ensemble.

Enfin, je tiens à adresser mes sincères remerciements à l'ensemble de mes Enseignants, qui, par leur savoir, leur rigueur et leur dévouement, ont contribué à ma formation et m'ont guidé sur le chemin de la connaissance.

À toutes ces personnes chères à mon cœur, je dédie ce travail avec reconnaissance et respect.

RIAHI Mohammed Salim
Tlemcen, le 2 Juin 2025

Dedicaces 2

JE dédie ce travail à toutes celles et tous ceux qui, de près ou de loin, ont contribué à faire de ce projet ce qu'il est aujourd'hui. Sans eux, rien de tout cela n'aurait été possible.

À mon père, cet homme exceptionnel, qui m'a transmis les valeurs essentielles de la vie. Son amour, son courage et son soutien inébranlable ont été mes plus grandes sources d'inspiration. Grâce à lui, j'ai appris la rigueur, la persévérance et l'humilité. Il est et restera mon modèle, ma référence.

À ma mère, celle qui a été ma complice, ma confidente et mon soutien sans faille. Elle a sacrifié tant de choses pour mon bonheur, toujours présente, toujours prête à m'encourager, à me protéger et à me pousser à aller de l'avant. Elle est la source de ma force intérieure et de ma résilience.

À mes sœurs, qui ont toujours été là pour moi, me guidant par leur sagesse et leur bienveillance. Leur amour et leurs conseils ont été un phare dans les moments d'incertitude. Elles ont cru en moi avant même que j'y croie moi-même. Leur soutien constant a été un moteur essentiel dans ma réussite.

À mes grands-parents, qui ont enrichi ma vie de leurs expériences et de leurs enseignements. Leur sagesse et leurs histoires m'ont permis de me construire. Que Dieu leur accorde une longue vie pleine de santé et de sérénité.

OUAHRANI Mohammed
Tlemcen, le 2 Juin 2025

Remerciements

Ce document présente les travaux effectués dans le cadre de notre projet de fin d'étude de Master au Département de Génie Électrique et Électronique de la Faculté de Technologie de l'Université Abou Bekr Belkaïd de Tlemcen.

En premier lieu, nous tenons à exprimer notre gratitude à notre encadrant, Dr. BOURI Sihem, Maître de Conférences à l'Université Abou Bekr Belkaïd de Tlemcen, pour nous avoir proposé et dirigé ce travail, nous la remercions pour nous avoir fait profiter de son expérience, pour orientations qui nous ont permis de mener à bien ce travail.

Nous tenons aussi à exprimer notre gratitude à notre co-encadrant, Pr. MESSAOUD Hamouda, professeur à l'université d'Adrar, pour son aide lors de notre stage, nous le remercions pour nous avoir fait profiter de son expérience qui nous a permis de mener à bien ce travail.

Nous sommes sensibles à l'honneur que nous a fait Dr. DRIS Younes, Maître de conférences à l'Université Abou Bekr Belkaïd, pour avoir accepté de présider et de nous honorer de sa présence au sein du jury de soutenance du présent mémoire ; qu'il trouve ici l'expression de notre reconnaissance et de notre respect.

Enfin, nous tenons également à adresser nos vifs remerciements au Dr. MECHER-NENE Abdelkader, Maître de Conférences à l'Université Abou Bekr Belkaïd, nous le remercions chaleureusement pour avoir accepté d'examiner le présent mémoire et pour son observations et remarques pertinentes et constructives.

Nous tenons aussi à adresser nos sincères remerciements à tous les enseignants du département génie électrique et électronique pour leurs enseignements et les cours intéressants qu'ils nous ont prodigués pendant notre cursus à l'Université Abou BekrBelkaid-Tlemcen.

Merci enfin à tous ceux qui, de près ou de loin, nous ont aidé et donc ont contribué au succès de ce travail.

OUAHRANI Mohammed El Amine

*RIAHY Mohammed Salim
Tlemcen, le 2 Juin 2025*

TABLE DES MATIÈRES

DEDICACES 1	i
DEDICACES 2	ii
REMERCIEMENTS	iii
NOMENCLATURE	ix
GLOSSAIRE	xii
INTRODUCTION GÉNÉRALE	1
I ETAT DE L'ART SUR LES SYSTÈMES PHOTOVOLTAÏQUES	3
I.1 INTRODUCTION	4
I.2 L'ÉNERGIE SOLAIRE	4
I.3 L'INSTALLATION PHOTOVOLTAÏQUE	4
I.3.1 Les avantages	6
I.3.2 Les inconvénients	6
I.4 TECHNOLOGIE DES CELLULES PHOTOVOLTAÏQUES	6
I.4.1 Silicium monocristallin	6
I.4.2 Silicium polycristallin	7
I.4.3 Silicium amorphe	8
I.5 SCHÉMA ÉQUIVALENT D'UNE CELLULE PHOTOVOLTAÏQUE	8
I.6 LE MODULE PHOTOVOLTAÏQUE	9
I.6.1 Les diodes en série	10
I.6.2 Les diodes en parallèle	10
I.6.3 Caractéristiques d'un module PV	11
I.7 CARACTÉRISTIQUES D'UNE CELLULE PHOTOVOLTAÏQUE	11
I.7.1 caractéristique courant–tension (I–V)	11
I.7.2 Caractéristique puissance–tension (P–V)	12
I.8 TYPES D'INSTALLATIONS PHOTOVOLTAÏQUES	12
I.8.1 Installation hors réseau	12
I.8.2 Installation raccordée au réseau	13
I.8.2-a Injection totale de l'énergie électrique	13
I.8.2-b Injection de surplus	14
I.9 CONCLUSION	14
II LES MÉTHODES DE PRÉDICTION DE LA PUISSANCE PHOTOVOLTAÏQUE	15

II.1	INTRODUCTION	16
II.2	LA MÉTHODE PHYSIQUE	16
II.2.1	Modèles Géométriques (détermination de Position du Soleil)	16
II.2.1-a	Paramètres de position	16
II.2.1-b	Calcul des Angles Solaires (Zénithal, Azimutal, d'élévation)	17
II.2.1-c	L'irradiance extraterrestre sur une surface horizontale	19
II.2.1-d	L'irradiance extraterrestre sur des surfaces inclinées	19
II.2.2	Estimation de l'Irradiance Solaire incliné	21
II.2.2-a	Irradiance Directe sur la Surface Inclinée (I_{β})	21
II.2.2-b	Irradiance Réfléchie par le Sol sur la Surface Inclinée (d_{re})	21
II.2.2-c	Irradiance Diffuse sur la Surface Inclinée (D_{β})	22
II.2.3	prédictions de puissance	23
II.3	LES MÉTHODES STATISTIQUES	23
II.3.1	Régression Linéaire Simple (RLS)	23
II.3.1-a	Estimation du modèle par la méthode des moindres carrés ordinaires (MCO)	24
II.3.1-b	Test de la pente de la droite de régression	25
II.3.2	Régression linéaire multiple	25
II.3.2-a	Modèle de la RLM	26
II.3.2-b	Notation matricielle	26
II.3.2-c	Estimation du modèle par la méthode des moindres carrés ordinaires (MCO)	27
II.3.2-d	Estimation de la variance des résidus	27
II.3.2-e	Décomposition de la variabilité	27
II.3.2-f	Coefficients de détermination et ajusté	28
II.3.2-g	Prévision	29
II.3.3	Filtre de kalman	30
II.3.3-a	Historique	30
II.3.3-b	Présentation du filtre de kalman standard	30
II.3.3-c	Fonctionnement du filtre de kalman	31
II.3.3-d	Les Hypothèses	33
II.3.3-e	Algorithme du Filtre de kalman standard	33
II.4	MÉTHODES D'INTELLIGENCE ARTIFICIELLE	34
II.4.1	Réseaux de neurones	34
II.4.1-a	Historique	35
II.4.1-b	Réseau de neurones : fondement biologique	36
II.4.1-c	Réseaux de neurones artificiels	36
II.4.1-d	Différents modèles de RNA	39
II.4.1-e	Méthodes d'apprentissage ou d'entraînement	40
II.4.1-f	Réseau de neurones monocouche	41
II.4.1-g	Lois d'apprentissage	41
II.4.1-h	Réseau de neurones multicouche : perceptron multicouches (MultiLayers Perceptron (MLP))	43
II.5	CONCLUSION	46

III SIMULATION ET RÉSULTATS	47
III.1 INTRODUCTION	48
III.2 DESCRIPTION TECHNIQUE DU SITE ÉTUDIÉ : CENTRALE PHOTOVOLTAÏQUE D'ADRAR	48
III.3 ANALYSE DES DONNÉES	49
III.3.1 Interprétation des figures	50
III.4 PRÉDICTION DE LA PUISSANCE	52
III.4.1 Application de la méthode RLM	54
III.4.1-a Détermination des coefficients de corrélation	54
III.4.1-b Résultats et interprétation	55
III.4.2 Application de réseau de neurone	57
III.4.2-a Description de structure de réseau de neurone	58
III.4.2-b Paramètres du réseau Elman	58
III.4.2-c Résultats et interprétation	59
III.4.3 Régression Linéaire Multiple (RLM) combinée avec Réseau Elman (NAR)	61
III.4.3-a Structuration de la méthode	61
III.4.3-b Structuration globale du modèle	61
III.4.3-c Résultats	62
III.4.4 Régression Linéaire Multiple (RLM) combinée avec Filtre de Kalman	64
III.4.4-a Méthodologie	64
III.4.4-b structurations de la méthode	66
III.4.4-c Résultats	67
III.4.5 Comparaison entre les 4 méthodes	68
III.5 CONCLUSION	70
CONCLUSION GÉNÉRALE	71

TABLE DES FIGURES

I.1	Installation photovoltaïque	5
I.2	principe de fonctionnement des panneaux [5]	5
I.3	cellules en silicium monocristallin [7]	7
I.4	cellules en silicium polycristallin [8]	7
I.5	Cellule au Silicium amorphe (couche mince) [9]	8
I.6	Rendement des différentes technologies [10]	8
I.7	Schéma équivalent d'une cellule PV couplé à un récepteur [11]	9
I.8	Module photovoltaïque [12]	10
I.9	Schéma synoptique d'un générateur PV [13]	10
I.10	Caractéristique courant – tension d'une cellule photovoltaïque [19]	11
I.11	Caractéristique puissance-tension [20]	12
I.12	Installation hors réseau [21]	13
I.13	Installation raccordée au réseau [22]	14
I.14	Installation avec injection de surplus [23]	14
II.1	coordonnées des Paramètres de position [28]	17
II.2	Le dispositif de mesure à différents plans d'inclinaison [31]	20
II.3	Exemple de liaisons affine entre x et y [42]	24
II.4	Exemple de liaisons non affine entre x et y [43]	25
II.5	Exemple pas de liaisons apparente entre x et y [43]	25
II.6	Explication géométrique de la décomposition de la variabilité.	28
II.7	Etapes du filtre de kalman	32
II.8	Processus itératif du filtre de kalman linéaire.	33
II.9	Neurone biologique [47]	36
II.10	Neurone biologique / neurone artificiel. [47]	37
II.11	Neurone Artificiel [48]	38
II.12	Fonctions de transfert fréquemment utilisées	39
II.13	IRéseaux directs (feedforward networks)	39
II.14	Réseaux récurrents (feedback networks).	40
II.15	Processus d'apprentissage supervisé.	40
II.16	Processus d'apprentissage non-supervisé.	41
II.17	structure de Perceptron de Rosenblatt (1958) [51]	41
II.18	Perceptron multicouches	43
II.19	Schéma explicatif de l'algorithme de rétropropagation du gradient	44
III.1	parc des panneaux solaire au cite d'Adrar	49
III.2	corrélation entre la température ambiante et la puissance	50

III.3	Corrélation entre l'irradiance et la puissance	50
III.4	corrélation entre la température de module et la puissance	51
III.5	Corrélations entre l'humidité et la puissance	51
III.6	La corrélation entre puissance de sortie PV et la puissance prédit	55
III.7	Représentation de la puissance de sortie PV et la puissance prédit (RLM)	56
III.8	la structure de réseau de neurone	58
III.9	Représentation de la puissance de sortie PV et la puissance prédit (NAR)	59
III.10	Structuration globale de RLM+RNN	62
III.11	Représentation comparaison entre la puissance de sortie PV et la puissance prédit (methode de RLM+NAR)	62
III.12	structuration de la méthode RLM + filtre de Kalman	66
III.13	Représentation de la puissance de sortie PV et la puissance prédit (methode de RLM+filtre de Kalman)	67

Nomenclature

Les principales notations et abréviations utilisées dans ce mémoire sont explicitées ci-dessous, sous leur forme la plus couramment employée dans le domaine du génie électrique.

Symbole	Signification
θ_z	Angle Zénithal
δ	La Déclinaison Solaire
ω	L'angle horaire
α	L'Angle d'élévation
γ_s	L'Angle Azimutal
φ	La latitude
G_o	L'irradiance extraterrestre sur une surface horizontale
G_{on}	L'irradiance extraterrestre normale
G_{sc}	La constante solaire
n	Le numéro du jour de l'année
β	Angle d'inclinaison
θ	Angle d'Incidence
G_β	L'Irradiance Solaire incliné
I_β	Irradiance Directe sur la Surface Inclinée
I_h	Irradiance sur la Surface direct horizontal
R_b	Coefficient de correction
d_{re}	Irradiance Réfléchie par le Sol sur la Surface Inclinée
ρ	Coefficient de réflexion de la lumière diffuse du sol
GH	Irradiance global horizontal
D_β	Irradiance Diffuse sur la Surface Inclinée
d_h	Irradiance Diffuse sur la Surface Inclinée horizontal
A_i	L'indice d'anisotropie
I_o	Irradiance global extraterrestre incident sur un plan horizontal
f	Paramètre qui tient compte de l'état du ciel
η	Rendement des panneaux
γ	Coefficient de température
γ_i	i^{me} observation de la variable Y : dépendante
X	Matrice des variables indépendantes (température, irradiation)
ϵ_i	Bruit
β	Coefficient de régression

H	Humidité
b	Biais de réseau de neurone
μ	Le taux d'apprentissage
$y_d(k)$	La réponse désirée ou attendue par le réseau
$x_i(k)$	L'entrée du neurone
SD	écart type de l'erreur
K	Gain de Kalman
T_a	Température ambiante
T_m	Température de module
I	Irradiation
Y_K	Vecteur de mesure
X_K	Vecteur d'état
A_k	Matrice de transition
B_k	Matrice de commande à l'instant K
C_k	Matrice d'observation (mesure)
U_k	Vecteur des commandes appliquées
W_k	Bruit de modélisation
V_k	Bruit de mesure
Q_k	Matrice de variances-covariances du bruit de processus à l'instant K
R_k	Matrice de variances-covariances du bruit de mesure à l'instant K
P_k	La covariance de l'erreur
e_k	erreur
$P_m(i)$	La puissance moyenne horaire
$P(t)$	La puissance instantanée mesuré à chaque 15min
N	Nombre de mesure à chaque heure
I	le nombre d'heures
$I_m(i)$	L'irradiation moyenne horaire
$T_m(i)$	Température moyenne horaire
P_{prdit}	Vecteur de puissance prédit
R^2	Coefficients de détermination
R^2_{adj}	Coefficients de détermination ajusté
f(s)	Fonction d'activation
w_i	Poids synoptiques
s	Vecteur synoptiques

Glossaire

Acronyme	Signification
NMBE	Erreur de biais moyenne normalisée
NMAE	Erreur absolue moyenne normalisée
NRMSE	Erreur quadratique moyenne normalisée
SCE_t	Somme des carrés totaux (centrés)
SCE_{reg}	Somme des carrés expliqués (centrés)
SCE_r	Somme des carrés des résidus
STC	Condition standard
TSV	Le temps solaire vrai
RLS	Régression Linéaire Simple
RLM	Régression linéaire multiple
RNN	Réseaux de neurones récurrents
RNA	Réseaux de neurone artificiel
NAR	Réseau autorégressif non linéaire
KF	Filtre de Kelman standard
EKF	Filtre de Kelman étendu
UKF	Filtre de Kelman non parfumé

Introduction générale

Face aux défis environnementaux et énergétiques croissants, la transition vers les énergies renouvelables s'impose aujourd'hui comme une nécessité à l'échelle mondiale. Parmi les différentes sources d'énergie propre, l'énergie solaire occupe une place stratégique en raison de son abondance, de son caractère inépuisable et de son impact environnemental réduit. L'Algérie, pays fortement ensoleillé, dispose d'un immense potentiel solaire, notamment dans les régions sahariennes où le taux d'ensoleillement dépasse les 3 000 heures par an.

Dans ce contexte, le développement de centrales solaires photovoltaïques constitue une solution pertinente pour diversifier le mix énergétique national et réduire la dépendance aux énergies fossiles. La centrale photovoltaïque d'Adrar, mise en service en 2015, s'inscrit dans cette dynamique. Équipée de 75 000 panneaux en silicium polycristallin, elle représente un modèle de déploiement de la technologie solaire dans des conditions climatiques extrêmes.

Cependant, la nature intermittente de la production solaire - fortement influencée par des variables climatiques telles que l'irradiation, la température ou encore l'humidité - rend la prédiction de l'énergie produite complexe. Une estimation fiable de cette production est pourtant essentielle pour la gestion du réseau électrique, la planification énergétique, ainsi que pour l'optimisation du fonctionnement et du dimensionnement des systèmes.

Ce mémoire s'intéresse ainsi à la prédiction de la production d'énergie photovoltaïque sur le site d'Adrar, en s'appuyant sur des approches statistiques et des techniques d'intelligence artificielle adaptées aux séries temporelles météorologiques et énergétiques. L'objectif principal est de proposer un modèle de prévision performant permettant d'anticiper la production d'électricité solaire à court terme, et ainsi contribuer à une meilleure intégration des énergies renouvelables dans le réseau.

Notre mémoire sera divisé comme suit :

- Chapitre 1 : On discutera sur les systèmes photovoltaïques
- Chapitre 2 : On parlera des différentes méthodes pour l'estimation de l'énergie électrique

- Chapitre 3 : On appliquera quelques méthodes pour l'estimation de la puissance d'un parc photovoltaïque situé à Adrar, et on fera une étude comparative entre ces méthodes
- On terminera par une conclusion générale

Chapitre I

Etat de l'art sur les systèmes photovoltaïques

SOMMAIRE

I.1	INTRODUCTION	4
I.2	L'ÉNERGIE SOLAIRE	4
I.3	L'INSTALLATION PHOTOVOLTAÏQUE	4
I.3.1	Les avantages	6
I.3.2	Les inconvénients	6
I.4	TECHNOLOGIE DES CELLULES PHOTOVOLTAÏQUES	6
I.4.1	Silicium monocristallin	6
I.4.2	Silicium polycristallin	7
I.4.3	Silicium amorphe	8
I.5	SCHÉMA ÉQUIVALENT D'UNE CELLULE PHOTOVOLTAÏQUE	8
I.6	LE MODULE PHOTOVOLTAÏQUE	9
I.6.1	Les diodes en série	10
I.6.2	Les diodes en parallèle	10
I.6.3	Caractéristiques d'un module PV	11
I.7	CARACTÉRISTIQUES D'UNE CELLULE PHOTOVOLTAÏQUE	11
I.7.1	caractéristique courant–tension (I–V)	11
I.7.2	Caractéristique puissance–tension (P–V)	12
I.8	TYPES D'INSTALLATIONS PHOTOVOLTAÏQUES	12
I.8.1	Installation hors réseau	12
I.8.2	Installation raccordée au réseau	13
I.9	CONCLUSION	14

I.1 Introduction

L'énergie photovoltaïque permet de convertir directement l'énergie lumineuse du soleil en électricité à l'aide de cellules solaires conçues à partir de matériaux semi-conducteurs, principalement le silicium. En reliant ces cellules entre elles, en série ou en parallèle, on obtient un module photovoltaïque. Pour atteindre une puissance adaptée aux besoins, plusieurs modules peuvent à leur tour être interconnectés pour constituer un générateur photovoltaïque (PV). La quantité d'énergie produite par ce générateur dépend notamment de l'intensité de l'ensoleillement et de la température de jonction des cellules [1][2].

Dans ce chapitre, nous allons présenter des notions générales sur le système photovoltaïque. Nous commencerons par expliquer le principe de fonctionnement d'une installation photovoltaïque, en montrant comment l'énergie solaire est convertie en électricité. Ensuite, nous aborderons les différentes technologies de cellules photovoltaïques, en distinguant les principales familles comme le silicium monocristallin, polycristallin et amorphe.

Le schéma équivalent d'une cellule photovoltaïque sera brièvement présenté pour illustrer son comportement électrique à travers un modèle simplifié. Nous introduirons ensuite le module photovoltaïque, en mettant en évidence le rôle des diodes de protection et ses principales caractéristiques.

Par la suite, nous évoquerons les caractéristiques électriques d'une cellule photovoltaïque, notamment les courbes courant-tension (I-V) et puissance-tension, qui permettent de comprendre son rendement. Enfin, ce chapitre se clôturera par une classification des types d'installations photovoltaïques, qu'elles soient autonomes ou connectées au réseau électrique.

I.2 L'énergie solaire

L'énergie solaire est issue du rayonnement émis par le soleil et peut être exploitée pour produire de l'électricité à l'aide de la technologie photovoltaïque. Cette conversion s'effectue grâce à des panneaux solaires qui transforment directement l'énergie solaire en énergie électrique, constituant ainsi une solution propre et renouvelable pour la production d'électricité.

I.3 L'installation photovoltaïque

Le principe de fonctionnement d'une installation photovoltaïque est relativement simple : il s'agit de convertir le rayonnement du soleil en électricité.

Cette transformation repose sur un phénomène physique connu sous le nom d'effet photovoltaïque.

Schématiquement, l'effet photovoltaïque se produit lorsqu'un photon est absorbé par un matériau constitué de semi-conducteurs dopés de type p (positif) et n (négatif), formant ce qu'on appelle une jonction p-n (ou n-p). En raison de ce dopage, un champ électrique



FIGURE I.1 – Installation photovoltaïque

permanent s'établit dans le matériau, de manière similaire à un champ magnétique permanent créé par un aimant.

Lorsqu'un photon incident (grain de lumière) entre en interaction avec les électrons d'un matériau semi-conducteur, il transfère son énergie à un électron situé dans la bande de valence. Cet électron est alors libéré et devient mobile, soumis au champ électrique intrinsèque présent dans la jonction. Sous l'effet de ce champ, l'électron migre vers la face supérieure du matériau, tandis qu'un "trou" - laissé par l'électron - se déplace en sens inverse. Des électrodes placées respectivement sur les faces supérieure et inférieure permettent de collecter les électrons et de les faire circuler à travers un circuit externe, où ils effectuent un travail électrique avant de rejoindre les trous sur la face opposée [3][4].

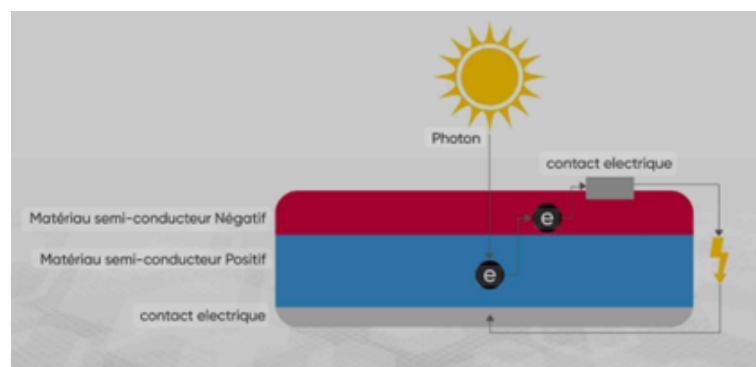


FIGURE I.2 – principe de fonctionnement des panneaux [5]

Une cellule photovoltaïque est composée d'un matériau semi-conducteur, le plus souvent du silicium, et conçue de manière à permettre la collecte des électrons libérés afin de générer un courant électrique. Pour obtenir un courant exploitable à l'échelle pratique, plusieurs cellules sont assemblées. Cet ensemble constitue ce qu'on appelle un module photovoltaïque, plus communément connu sous le nom de panneau solaire [3].

I.3.1 Les avantages

L'énergie électrique produite est propre à l'utilisation, sans émissions polluantes, ce qui en fait une solution respectueuse de l'environnement, conforme aux principes du développement durable.

Il s'agit d'une source d'énergie renouvelable, car le rayonnement solaire est inépuisable à l'échelle humaine.

I.3.2 Les inconvénients

Le coût des systèmes photovoltaïques reste relativement élevé, car ils proviennent d'une technologie de pointe. Le coût total est fortement influencé par la puissance de crête installée.

Le rendement des cellules photovoltaïques disponibles pour le grand public demeure assez faible, avoisinant les 10 %, ce qui limite la puissance réellement délivrée.

Le marché du photovoltaïque reste encore restreint, bien qu'il soit en pleine expansion.

La production d'électricité est limitée aux heures d'ensoleillement, alors que la demande maximale, notamment chez les particuliers, survient généralement la nuit. Le stockage de l'énergie produite représente un véritable défi, les solutions actuelles (notamment les batteries) étant encore coûteuses et à fort impact écologique.

I.4 Technologie des cellules photovoltaïques

Il existe plusieurs types de cellules solaires photovoltaïques, chacune présentant des caractéristiques spécifiques en termes de rendement et de coût. Néanmoins, quel que soit le type de cellule utilisé, le rendement global reste relativement limité, variant généralement entre 5 % et 22 % de l'énergie solaire reçue.

Les cellules photovoltaïques sont classées en trois grandes catégories, en fonction du matériau de base utilisé [6] :

- Cellules monocristallines
- Cellules polycristallines
- Cellules à couches minces

I.4.1 Silicium monocristallin

Pour ce genre d'applications technologiques, le silicium pur est obtenu à partir de la silice de quartz ou de sable par transformation chimique métallurgique. Le silicium a un rendement électrique et une durée de vie de l'ordre de deux fois celle du silicium amorphe, mais il est nettement plus cher [6].



FIGURE I.3 – cellules en silicium monocristallin [7]

I.4.2 Silicium polycristallin

Le silicium polycristallin est un matériau constitué de plusieurs cristaux de silicium juxtaposés, formés généralement par un procédé de moulage. Ce type de silicium est moins coûteux à produire que le silicium monocristallin, ce qui en fait une solution économiquement plus accessible.

Les cellules fabriquées à partir de ce matériau ont généralement une forme carrée ou rectangulaire, ce qui facilite leur intégration et leur assemblage dans les modules photovoltaïques.

Les cellules polycristallines sont caractérisées par [6] :

- Coût de production moins élevée.
- Nécessite moins d'énergie.
- Rendement typique de 11 % et jusqu'à 13 % en labo.

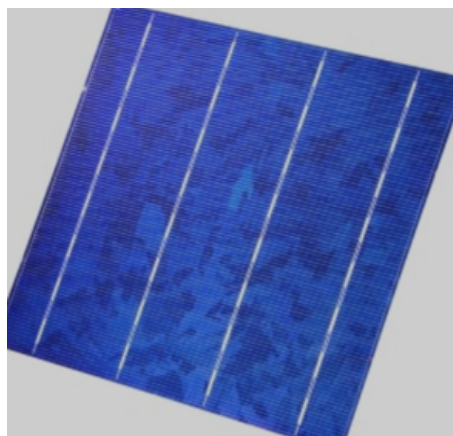


FIGURE I.4 – cellules en silicium polycristallin [8]

I.4.3 Silicium amorphe

Le silicium amorphe a la particularité d'absorber le rayonnement solaire jusqu'à cent fois plus efficacement que le silicium à l'état cristallin. Cette propriété permet la fabrication de cellules à partir de couches extrêmement minces, ce qui réduit considérablement la quantité de matériau nécessaire et permet une plus grande souplesse dans les applications (souvent sur des surfaces souples ou intégrées à des objets) [6].

Les cellules à couches minces sont caractérisées par :

- Coût de production plus bas.
- Rendement de seulement 5 % par module et 10 % au laboratoire.
- Fonctionne sous de très faible éclairement.



FIGURE I.5 – Cellule au Silicium amorphe (couche mince) [9]

Technologie	Rendement typique (%)	Rendement théorique (%)
Monocristallin	12-16	24
Polycristallin	11-13	18,6
Amorphe	5-10	12,7

FIGURE I.6 – Rendement des différentes technologies [10]

I.5 Schéma équivalent d'une cellule photovoltaïque

Le schéma équivalent d'une cellule photovoltaïque inclut un générateur de courant, représentant l'éclairage solaire, ainsi qu'une diode montée en parallèle, qui modélise la jonction P-N. Toutefois, pour refléter plus fidèlement le comportement réel de la cellule, le modèle intègre également les effets des résistances parasites liées au processus de fabrication. Ces pertes sont représentées dans le schéma équivalent par deux résistances : une en série et une en parallèle [11].

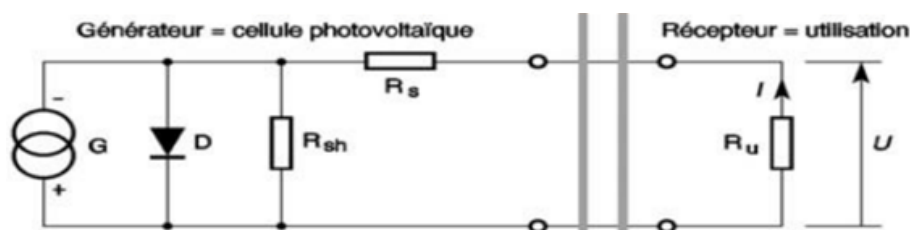


FIGURE I.7 – Schéma équivalent d'une cellule PV couplé à un récepteur [11]

G : est une source de courant parfaite.

D : est une diode matérialisant le fait que le courant ne circule que dans un sens.

R_{sh} : est une résistance shunt qui prend en compte les fuites inévitables de courant qui interviennent entre les bornes opposées positive et négative d'une photopile (micro court-circuit dans le silicium en particulier).

R_s : est une résistance série qui est due aux différentes résistances électriques que le courant rencontre sur son parcours (résistance intrinsèque des couches, résistance des contacts).

R_u : est l'impédance du récepteur qui impose le point de fonctionnement sur la photopile en fonction de sa caractéristique courant tension à l'éclairement considéré (dans le cas où le récepteur est assimilable à un résistor).

I.6 Le module photovoltaïque

Pour augmenter la puissance produite, plusieurs cellules solaires sont assemblées afin de former un module, comme illustré dans la (Figure I.8) Les connexions en série de ces cellules permettent d'augmenter la tension tout en maintenant le même courant, tandis que la mise en parallèle permet d'augmenter le courant tout en conservant la même tension. Afin de protéger les cellules de l'humidité, elles sont encapsulées dans un polymère EVA (éthylène-vinyl-acétate). De plus, la surface avant des modules est protégée par un verre trempé à haute transmission et présentant une bonne résistance mécanique, tandis que la surface arrière est généralement protégée par une couche de polyéthylène [12].

Actuellement, la puissance d'un module photovoltaïque varie de quelques watts-crête à quelques dizaines de watts-crête. Pour atteindre des puissances plus élevées, plusieurs modules sont associés en série et en parallèle afin de constituer un générateur photovoltaïque, comme illustré à la figure (I.10). Cependant, en cas de déséquilibres importants causés par l'occultation partielle ou totale de certaines cellules, les modules peuvent fonctionner comme des récepteurs, ce qui entraîne des pertes de performance. Pour pallier ces effets, des diodes de protection sont généralement insérées en parallèle ou en série avec les modules [13,14, 15, 16].



FIGURE I.8 – Module photovoltaïque [12]

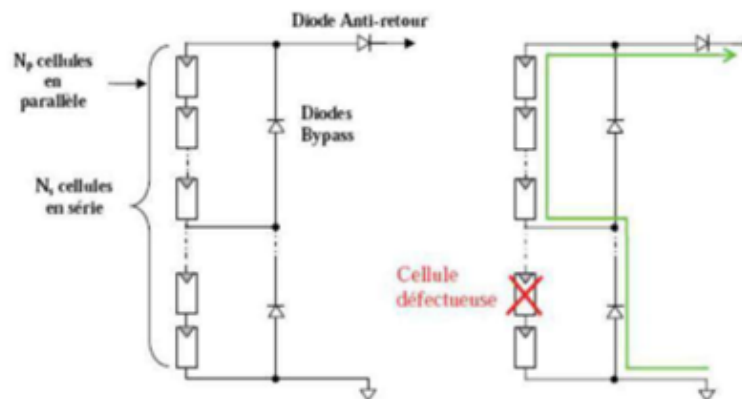


FIGURE I.9 – Schéma synoptique d'un générateur PV [13]

I.6.1 Les diodes en série

Pour empêcher la batterie de se décharger la nuit dans les cellules PV ou pour empêcher une série de modules contenant un module défaillant ou masqué de devenir réceptrice du courant fourni par les autres séries, qui ont alors une tension plus élevée qu'elle, une diode est intégrée dans chaque série. On l'appelle aussi diode anti-retour et elle est située en série avec les modules.

I.6.2 Les diodes en parallèle

Lorsqu'un module d'une chaîne série ne peut plus produire d'énergie - en raison d'un ombrage, d'un encrassement ou d'une défaillance - il risque de se comporter comme une charge réceptrice, ce qui peut entraîner sa détérioration. Pour éviter ce phénomène, des diodes dites de "by-pass" sont placées en parallèle avec chaque module du générateur. Ces diodes permettent de dériver le courant généré par les autres modules actifs, évitant ainsi que le module défaillant subisse une surcharge ou une surchauffe [14, 17, 18].

I.6.3 Caractéristiques d'un module PV

- La puissance de crête, P_c : Puissance électrique maximum que peut fournir le module dans les conditions standards (25°C et un éclairement de $1000\text{W}/\text{m}^2$). [17]
- La caractéristique I/V : Courbe représentant le courant I débité par le module en fonction de la tension aux bornes de celui-ci.
- Courant de court-circuit, I_{cc} : Courant débité par un module en court-circuit pour un éclairement " plein soleil ".
- Point de fonctionnement optimum, (U_m, I_m) : Lorsque la puissance de crête est maximum en plein soleil, $P_m = U_m \times I_m$.
- Rendement : Rapport de la puissance électrique optimale à la puissance de radiation incidente.
- Facteur de forme : Rapport entre la puissance optimale P_m et la puissance maximale que peut avoir la cellule : $V_{oc} \times I_{cc}$.

I.7 Caractéristiques d'une cellule photovoltaïque

L'analyse des performances d'une cellule photovoltaïque repose essentiellement sur l'étude de ces caractéristiques

I.7.1 caractéristique courant-tension (I-V)

Cette courbe permet de visualiser le comportement électrique de la cellule en fonction des conditions d'ensoleillement et de température, et constitue un outil fondamental pour évaluer son rendement et sa puissance maximale.

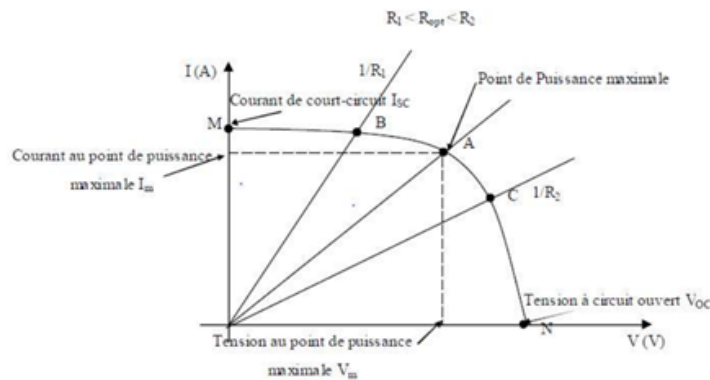


FIGURE I.10 – Caractéristique courant – tension d'une cellule photovoltaïque [19]

Comme illustré à la (Figure I.9), la cellule photovoltaïque présente une courbe courant-tension (I-V) non linéaire, fortement influencée par les conditions d'ensoleillement et de température. Cette courbe met en évidence trois points caractéristiques essentiels [19] :

- Le point de puissance maximale (PPM) : Il correspond au point où la cellule délivre la puissance électrique maximale.
- Le courant de court-circuit (I_{sc}) : C'est le courant maximal fourni par la cellule lorsque ses bornes sont en court-circuit (tension nulle).

- La tension de circuit ouvert (V_{oc}) : Il s'agit de la tension maximale mesurée aux bornes de la cellule lorsque le courant est nul (circuit ouvert), typiquement d'environ 0,6 V pour une cellule standard.

I.7.2 Caractéristique puissance–tension (P–V)

La courbe P–V d'une cellule photovoltaïque décrit l'évolution de la puissance délivrée en fonction de la tension à ses bornes. La puissance maximale atteinte est appelée puissance crête, notée W_c (Watt-crête) ou W_p (Watt-peak). Elle correspond à la puissance fournie par la cellule dans des conditions dites standards, définies comme suit [19] :

- un éclairement solaire de $1kW/m^2$,
- une température de cellule de $+25\text{ }^\circ\text{C}$.

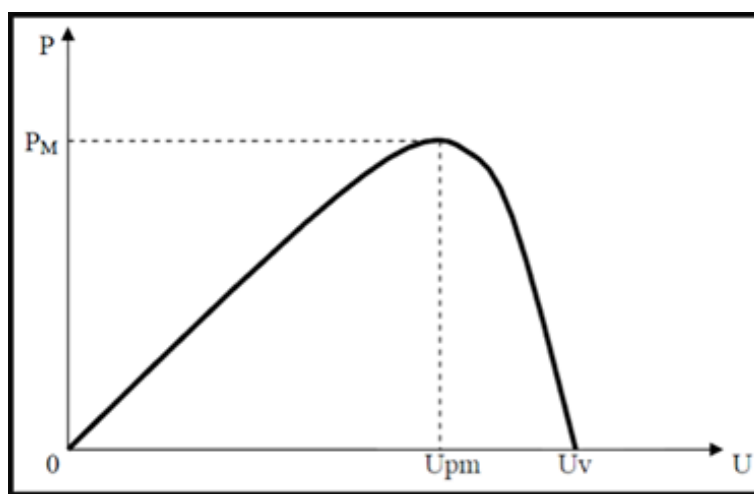


FIGURE I.11 – Caractéristique puissance-tension [20]

I.8 Types d'installations photovoltaïques

I.8.1 Installation hors réseau

Ce type d'installation, également appelée installation autonome, est conçu pour les sites isolés ne pouvant pas être raccordés au réseau électrique. L'énergie produite par les panneaux solaires est soit consommée immédiatement, soit stockée dans des batteries pour une utilisation différée, notamment durant la nuit ou les périodes sans ensoleillement. Ce système est souvent utilisé dans les zones rurales, les refuges, les stations de télécommunications ou les équipements de signalisation.

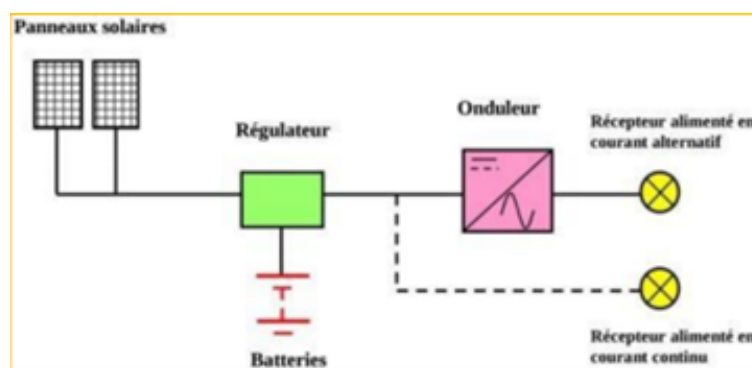


FIGURE I.12 – Installation hors réseau [21]

Pour répondre à la totalité des besoins en énergie dans une installation hors réseau, plusieurs composants sont nécessaires :

- Des accumulateurs (batteries) assurent le stockage de l'électricité afin de garantir une alimentation continue, y compris la nuit ou lors de mauvais temps.
- Les panneaux photovoltaïques produisent un courant continu à partir de l'énergie solaire.
- Le régulateur de charge gère la charge et la décharge des batteries en fonction de leur capacité, tout en assurant leur protection contre la surcharge ou la décharge profonde.
- L'onduleur convertit le courant continu issu des batteries ou des panneaux en courant alternatif, adapté aux récepteurs et appareils domestiques (AC).

I.8.2 Installation raccordée au réseau

I.8.2-a Injection totale de l'énergie électrique

Dans ce mode de fonctionnement, l'intégralité de l'énergie produite par les panneaux photovoltaïques est injectée dans le réseau public, sans être utilisée directement par le producteur. Elle est donc entièrement revendue au fournisseur d'électricité.

Ce type d'installation implique un double point de raccordement au réseau :

- Le premier point est le raccordement classique du consommateur, équipé d'un compteur standard pour sa propre consommation. Il ne bénéficie pas directement de sa production solaire.
- Le second point correspond à un nouveau branchement dédié à l'injection de la production solaire dans le réseau. Ce système inclut :
 - Un compteur de production, qui mesure l'énergie injectée
 - Un compteur de non-consommation, servant à vérifier qu'aucune consommation frauduleuse n'est réalisée à partir de cette production.

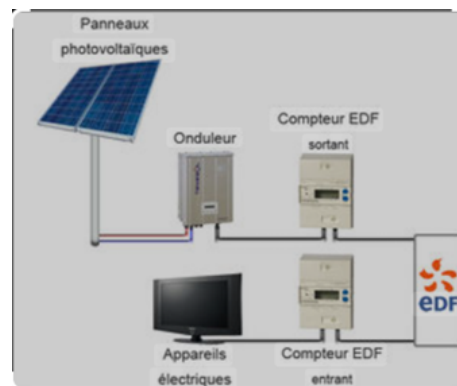


FIGURE I.13 – Installation raccordée au réseau [22]

I.8.2-b Injection de surplus

Cette solution repose sur un raccordement au réseau public en un seul point. L'utilisateur consomme en priorité l'énergie produite par son installation photovoltaïque. En cas d'excédent de production, le surplus est injecté dans le réseau électrique.

Inversement, lorsque la production solaire est insuffisante pour couvrir les besoins, le réseau complète automatiquement l'alimentation.

Dans ce cas, un seul compteur supplémentaire est ajouté à celui déjà existant pour mesurer l'énergie réinjectée.

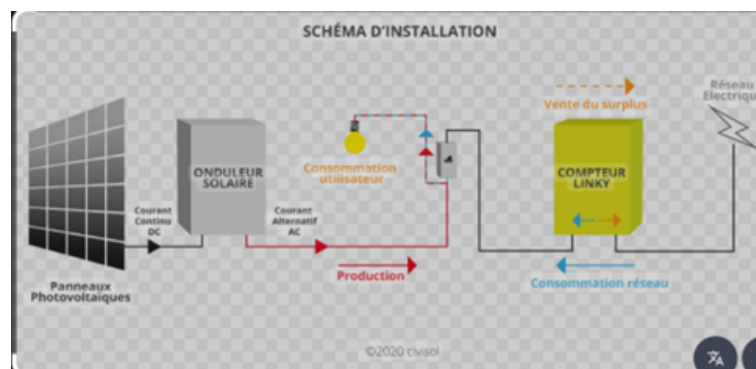


FIGURE I.14 – Installation avec injection de surplus [23]

I.9 Conclusion

Dans ce chapitre, nous avons présenté les bases du système solaire photovoltaïque, en expliquant le principe de conversion de l'énergie solaire en électricité grâce à l'effet photovoltaïque. Nous avons étudié les différents types de cellules (monocristallines, polycristallines, amorphes), leurs rendements et caractéristiques, ainsi que le fonctionnement des modules photovoltaïques et leurs composants (diodes). Enfin, nous avons distingué les installations hors réseau et celles raccordées au réseau électrique.

Chapitre II

Les méthodes de prédiction de la puissance photovoltaïque

SOMMAIRE

II.1	INTRODUCTION	16
II.2	LA MÉTHODE PHYSIQUE	16
II.2.1	Modèles Géométriques (détermination de Position du Soleil)	16
II.2.2	Estimation de l'Irradiance Solaire incliné	21
II.2.3	prédictions de puissance	23
II.3	LES MÉTHODES STATISTIQUES	23
II.3.1	Régression Linéaire Simple (RLS)	23
II.3.2	Régression linéaire multiple	25
II.3.3	Filtre de kalman	30
II.4	MÉTHODES D'INTELLIGENCE ARTIFICIELLE	34
II.4.1	Réseaux de neurones	34
II.5	CONCLUSION	46

II.1 Introduction

La prédiction de la puissance photovoltaïque joue un rôle essentiel dans l'intégration des énergies renouvelables au sein des réseaux électriques. Elle permet d'anticiper la production d'énergie, d'améliorer la gestion des systèmes photovoltaïques et de garantir une meilleure stabilité du réseau [24]. Cependant, cette tâche reste difficile en raison de la dépendance de la production photovoltaïque aux conditions météorologiques variables telles que l'irradiation, la température et la nébulosité [25].

Pour relever ce défi, deux grandes familles de méthodes sont généralement utilisées : les méthodes physiques et les méthodes statistiques. Les méthodes physiques s'appuient sur la modélisation des phénomènes solaires et atmosphériques, tandis que les méthodes statistiques exploitent les données historiques pour établir des relations empiriques entre les paramètres météorologiques et la puissance produite [26][27].

Ce chapitre présente les différentes méthodes de prédiction de la puissance photovoltaïque : d'une part une approche statistique reposant sur la régression linéaire multiple, les réseaux de neurones et le filtre de Kalman, et d'autre part une méthode physique fondée sur l'exploitation des paramètres météorologiques mesurés

II.2 La méthode physique

La méthode physique repose sur la modélisation des phénomènes liés au rayonnement solaire en utilisant des équations déterministes basées sur des paramètres météorologiques mesurés ou estimés.

II.2.1 Modèles Géométriques (détermination de Position du Soleil)

L'énergie solaire atteignant une surface donnée dépend directement de l'orientation de celle-ci et de la position du soleil. Pour récupérer un maximum d'énergie en provenance du soleil, il est nécessaire d'orienter au mieux le récepteur par rapport aux rayons lumineux. La connaissance de la position du soleil en fonction du temps est donc fondamentale.

II.2.1-a Paramètres de position

- L'équateur est une ligne fictive de séparation entre l'hémisphère nord et l'hémisphère sud, définie par un plan perpendiculaire à l'axe des pôles.
- Le méridien est le plan du lieu passant par l'axe des pôles.
- La latitude donne la localisation d'un point par rapport à l'équateur, elle varie de 0° à l'équateur à 90° N (ou S) aux pôles. Elle représente aussi la distance angulaire du site du soleil (S) par rapport au plan de l'équateur. Cette convention de signe affecte le signe (+) à tous les lieux de l'hémisphère nord et le signe (-) à tous les lieux de l'hémisphère sud.
- La longitude : C'est l'angle formé par le méridien de Greenwich et le méridien du lieu considéré. La longitude est comprise entre -180 (vers l'ouest) et $+180$ (vers l'est).

Comme la terre met 24 heures pour faire un tour sur elle-même (360°), chaque heure représente 15° d'écarte de longitude et donc, chaque degré de longitude représente 4 minutes [28].

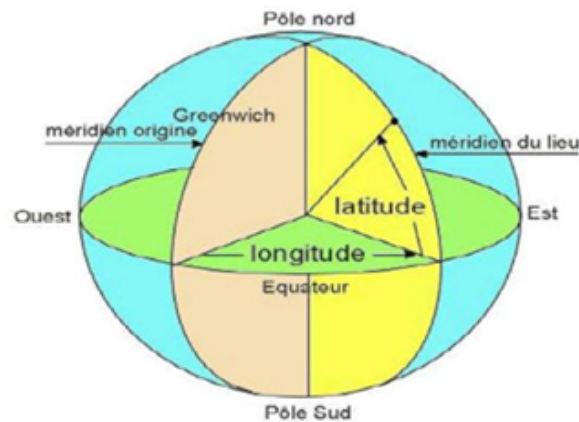


FIGURE II.1 – coordonnées des Paramètres de position [28]

II.2.1-b Calcul des Angles Solaires (Zénithal, Azimutal, d'élévation)

La détermination de la position du soleil est cruciale pour l'analyse des systèmes solaires. Elle s'effectue par le calcul des angles solaires : zénithal (θ_z), d'élévation (α) et azimutal (γ_s). Ces angles varient en fonction de la position géographique (latitude φ et longitude λ), du moment dans l'année (jour de l'année n) et de l'heure, en tenant compte des corrections temporelles et géographiques.

Le but de calculer l'angle zénithal et l'angle azimutal dans cette étude sur la centrale solaire est :

- Déterminer la position précise du soleil dans le ciel à tout moment donné. L'angle zénithal indique la hauteur du soleil, et l'angle azimutal indique sa direction horizontale.
- évaluer l'angle d'incidence du rayonnement solaire direct sur les panneaux photovoltaïques. Connaître la position du soleil permet de calculer l'angle sous lequel les rayons solaires frappent la surface des panneaux, ce qui influence directement la quantité d'énergie captée. Un angle d'incidence optimal (proche de 0°) maximise l'absorption.
- Optimiser l'orientation des panneaux solaires. L'angle azimutal du soleil est crucial pour déterminer l'orientation horizontale idéale des panneaux (généralement plein sud dans l'hémisphère nord) afin de maximiser l'exposition au soleil tout au long de la journée et de l'année.
- Analyser les ombrages. La trajectoire du soleil, décrite par les angles zénithal et azimutal au cours de la journée et des saisons, permet d'identifier et d'évaluer les ombrages potentiels causés par le relief, la végétation ou les structures environnantes, qui peuvent réduire la production d'énergie de la centrale.

1. Calcul de l'Angle Zénithal (θ_z) :

L'angle zénithal, l'angle entre la verticale et la direction du soleil, est donné par [28] :

$$\cos(\theta_z) = \sin(\varphi) \sin(\delta) + \cos(\varphi) \cos(\delta) \cos(\omega) \quad (\text{II.1})$$

$$\theta_z = \cos^{-1} [\sin \delta \cdot \sin \varphi + \cos \delta \cdot \cos \varphi \cdot \cos \omega] \quad (\text{II.2})$$

Où :

φ est la latitude du lieu (positive pour l'hémisphère nord, négative pour l'hémisphère sud).

δ : est la Déclinaison Solaire c'est l'angle entre le plan équatorial de la Terre et la direction du Soleil, est donné par [28] :

$$\delta = 23.45 \cdot \sin \left(360 \cdot \frac{n - 81}{365} \right) \quad (\text{II.3})$$

Où n représente le numéro du jour de l'année.

ω c'est L'angle horaire, qui mesure le déplacement angulaire du soleil par rapport au midi solaire local, est approximé par [28] :

$$\omega = 15 \cdot (TSV - 12) \quad (\text{II.4})$$

Où TSV est le temps solaire vrai, en un instant et un lieu donné, est l'angle horaire du soleil ω . Il est donné sous la forme suivante :

Si $\omega = 0$, TSV = 12h.00

Si $\omega < 0$, la matinée

Si $\omega > 0$, l'après-midi.

2. Calcul de l'Angle d'élévation (α)

C'est l'angle entre la droite joignant le centre du disque solaire au point d'observation et le plan horizontal passant par le point d'observation. La hauteur du soleil est bien représentée par une formule suivante [28] :

$$\sin \alpha = \cos \varphi \cdot \cos \delta \cdot \cos \omega + \sin \varphi \cdot \sin \delta \quad (\text{II.5})$$

$$\alpha = 90 - \theta_z = \sin^{-1} (\cos \varphi \cdot \cos \delta \cdot \cos \omega + \sin \varphi \cdot \sin \delta) \quad (\text{II.6})$$

3. Calcul de l'Angle Azimutal (γ_s)

L'angle azimutal solaire γ_s peut avoir une valeur comprise entre 180 et -180. Ce dernier est négatif si l'angle horaire est négatif et positif si l'angle horaire est positif. La fonction « sin » dans l'équation (1.7) est égale à +1 si ω est positif et est égale à -1 si ω est négatif [29] et [23] :

$$\gamma_s = \sin \omega \left| \cos^{-1} \left(\frac{\cos \theta_z \sin \varphi - \sin \delta}{\sin \theta_z \cos \varphi} \right) \right| \quad (\text{II.7})$$

$$\gamma_s = \begin{cases} + \cos^{-1} \left(\frac{\cos \theta_z \sin \varphi - \sin \delta}{\sin \theta_z \cos \varphi} \right); & \text{if } \omega > 0, \\ - \cos^{-1} \left(\frac{\cos \theta_z \sin \varphi - \sin \delta}{\sin \theta_z \cos \varphi} \right); & \text{if } \omega < 0. \end{cases} \quad (\text{II.8})$$

II.2.1-c L'irradiance extraterrestre sur une surface horizontale

L'irradiance extraterrestre sur une surface horizontale, notée G_o , correspond au flux solaire incident par unité de surface mesuré à la frontière de l'atmosphère, projeté sur un plan horizontal. Elle se déduit de l'irradiance extraterrestre normale G_{on} en la multipliant par le cosinus de l'angle zénithal solaire θ_z :

$$G_o = G_{on} \cos \theta_z \quad (\text{II.9})$$

Et G_{on} c'est l'irradiance extraterrestre normale Représente la puissance solaire moyenne reçue par unité de surface à une distance d'une unité astronomique (la distance moyenne entre la Terre et le Soleil), en l'absence d'atmosphère terrestre. [30]

$$G_{on} = G_{sc} \left[1 + 0.033 \cos \left(\frac{360}{365n} \right) \right] (W/m^2) \quad (\text{II.10})$$

Avec :

G_{sc} est la constante solaire, dont la valeur standard est d'environ $1367W/m^2$.

n : est le numéro du jour de l'année (allant de 1 pour le 1er janvier à 365 ou 366 pour la fin de l'année).

II.2.1-d L'irradiance extraterrestre sur des surfaces inclinées

Pour transformer l'irradiance horizontale sur des surfaces inclinées on tente en compte de l'orientation de la Surface Inclinée (Module Solaire) qui définit par deux angles qui sont :

- Angle d'inclinaison (β) c'est L'angle entre la surface et l'horizontale ($0^\circ \leq \beta \leq 90^\circ$).
- Angle azimutal de la surface (γ) c'est L'angle horizontal entre le sud vrai et la projection normale de la surface sur l'horizon (positif à l'ouest, négatif à l'est). Pour une surface orientée plein sud, $\gamma = 0^\circ$.

1. Angles β optimal mensuel et annuel

L'angle optimal mensuel est une inclinaison moyenne propre pour chaque mois, qui fait que la surface du capteur reçoit un maximum d'énergie durant la période considérée. Il peut être tiré une fois localiser le maximum de rayonnement, en faisant varier l'angle d'inclinaison (β) de 0° à 90° . Le même principe pour l'angle optimal annuel, avec une inclinaison fixe où on peut récupérer le maximum d'énergie au cours de l'année. [31]

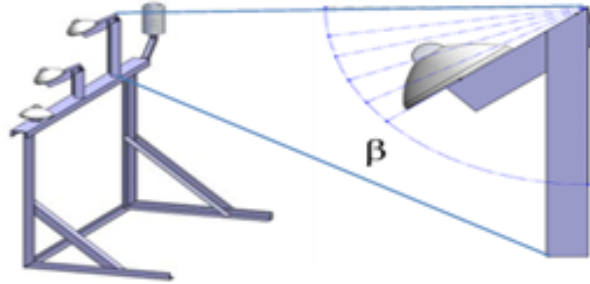


FIGURE II.2 – Le dispositif de mesure à différents plans d'inclinaison [31]

Le principe étant le calcul de l'énergie reçue sur une surface réceptrice d'orientation supposée fixe plane sud et une inclinaison variable entre 0° (plan horizontal) à un angle de 90° (plan vertical) pour chaque jour de l'année, de déduire par la suite l'irradiation moyenne mensuelle [31].

2. L'Angle d'Incidence (θ) :

L'angle d'incidence est l'angle sous lequel un rayon de lumière ou de rayonnement frappe une surface, mesuré à partir de la normale à la surface [23].

$$\theta = \cos^{-1} (\sin \varphi \cdot \sin \delta \cdot \cos \beta + \cos \delta \cdot \cos \gamma \cdot \cos \omega \cdot \sin \beta) + \cos \varphi (\cos \delta \cdot \cos \omega \cdot \cos \beta - \sin \delta \cdot \cos \gamma \cdot \sin \beta) + \cos \delta \cdot \sin \gamma \cdot \sin \omega \cdot \sin \beta$$

Tel que :

θ : Angle d'incidence

δ : Angle de déclinaison

ω : horaire

φ : Angle de latitude

β : Angle d'inclinaison

γ : Angle d'azimut de surface

II.2.2 Estimation de l'Irradiance Solaire incliné

Afin d'estimer avec précision l'irradiance captée par les modules photovoltaïques de la centrale, qui sont installés avec un angle d'inclinaison (β) et potentiellement un angle azimutal (γ), il est nécessaire de transformer les composantes de l'irradiance solaire (directe, diffuse et réfléchie) mesurées ou estimées sur une surface horizontale vers le plan incliné des modules. L'irradiance totale sur la surface inclinée (G_β) est la somme de l'irradiance directe (I_β), diffuse (D_β), et réfléchie par le sol (d_{re}) [33].

$$G_\beta = I_\beta + D_\beta + d_{re} \quad (\text{II.11})$$

II.2.2-a Irradiance Directe sur la Surface Inclinée (I_β)

L'irradiance directe sur la surface inclinée est obtenue en projetant l'Irradiance Directe Normale selon l'angle d'incidence (θ) [34] :

La composante directe est donnée par la relation suivante :

$$I_\beta = R_b \times I_h R_b = \frac{\cos \theta}{\cos \theta_z} \quad (\text{II.12})$$

I_h : éclaircissements direct horizontal.

R_b : c'est le coefficient de correction

Avec :

$$\cos \theta_z = \sin \varphi \sin \delta + \cos \varphi \cos \delta \cos \omega$$

Et

$$\begin{aligned} \cos \theta &= \sin \varphi (\sin \delta \cos \beta + \cos \delta \cos \gamma \cos \omega \sin \beta) + \\ &\cos \varphi (\cos \delta \cos \omega \cos \beta - \sin \delta \cos \gamma \sin \beta) + \cos \delta \sin \gamma \sin \omega \sin \beta \end{aligned}$$

II.2.2-b Irradiance Réfléchie par le Sol sur la Surface Inclinée (d_{re})

La surface inclinée reçoit également une contribution du Irradiance global réfléchi par le sol [30] :

$$d_{re} = \rho \frac{GH(1 - \cos \beta)}{2} \quad (\text{II.13})$$

Où :

ρ : représente le coefficient de réflexion de la lumière diffuse du sol et β , l'inclinaison du champ PV.

GH : C'est Irradiance global horizontal

II.2.2-c Irradiance Diffuse sur la Surface Inclinée (D_β)

La transformation de l'irradiance diffuse horizontale (d_h) vers une surface inclinée est plus complexe et plusieurs modèles existent, tenant compte de l'anisotropie du ciel :

1. Modèle Isotrope de Liu & Jordan :

Ce modèle suppose que le rayonnement diffus du ciel est isotrope, La relation de Liu & Jordan généralisée pour le calcul de l'éclairement Diffuse sur un plan incliné d'un angle (β) est donnée sous la forme suivante [35] :

$$D_\beta = d_h \left(\frac{1 + \cos \beta}{2} \right) \quad (\text{II.14})$$

L'éclairement global sur un plan incliné est donné par :

2. Modèle Anisotrope de Hay & Davies :

La relation générale pour le calcul du rayonnement global sur un plan incliné d'un angle (β) reste valable. Mais ce modèle suppose que le rayonnement diffus du ciel est formé de deux composantes, à savoir l'une isotrope et l'autre circumsolaire. Ce qui conduit à [36] :

$$D_\beta = d_h(1 - A_i) \left(\frac{1 + \cos \beta}{2} \right) + d_h(A_i) \left(\frac{\cos \theta}{\cos \theta_z} \right) \quad (\text{II.15})$$

A_i : Est l'indice d'anisotropie, il est donné par :

$$A_i = \frac{I_h}{I_o}$$

Avec :

I_h : L'éclairement direct incident sur un plan horizontal.

I_o : L'éclairement global extraterrestre incident sur un plan horizontal.

3. Modèle de Klucher :

Klucher a modifié le modèle de Temps and Coulson [34], il est donné comme suit :

$$D_\beta = d_h \left[\left(\frac{1 + \cos \beta}{2} \right) \left(1 + f \sin^3 \left(\frac{\beta}{2} \right) \right) (1 + f \cos^2 \theta \sin^3 \theta_z) \right] \quad (\text{II.16})$$

f : est un paramètre qui tient compte de l'état du ciel, et il se calcule par la relation suivante :

$$f = 1 - \left(\frac{d_h}{G_h} \right)^2$$

II.2.3 prédictions de puissance

Selon l'estimation de rayonnement globale incliné par le model de de Liu & Jordan :

$$G_{\beta} = I_{\beta} + d_h \left(\frac{1 + \cos \beta}{2} \right) + \rho \left(\frac{1 - \cos \beta}{2} \right) \quad (\text{II.17})$$

On peut estimer le la puissance d'un parc solaire à partir de de cette équation [27] :

$$P = G_{\beta} \cdot S \cdot \eta(T) \quad (\text{II.18})$$

Avec :

S : surface totale des panneaux

$\eta(T)$: rendement des panneaux dépendant de la température [37]

$$\eta(T) = \eta_{STC} \cdot [1 + \gamma(T - T_{STC})] \quad (\text{II.19})$$

Où :

η_{STC} est le rendement dans les condition stadard

$T_{STC} = 25$: la température nominale

γ : est le coefficient de température

II.3 Les méthodes statistiques

La méthode statistique consiste à exploiter les données historiques pour établir des relations mathématiques entre les variables météorologiques et la puissance photovoltaïque produite. Dans cette partie on s'intéresse sur quelque méthodes (régression linéaire multiple et simple et réseau de neurone et filtre de kalman)

II.3.1 Régression Linéaire Simple (RLS)

La régression linéaire est une méthode statistique utilisée pour analyser des données numériques [20]. Elle permet d'étudier la relation entre deux variables lorsqu'on dispose de données issues d'observations ou d'expériences.

Dans ce type d'analyse, une variable dépend d'une autre. En régression linéaire simple (RLS), il y a une seule variable qui explique (variable explicative) et une autre qui est influencée (variable expliquée). Cette relation peut être représentée par une fonction mathématique.

Pour une présentation assez complète du sujet, on renvoie aux ouvrages [38, 39, 40, 41] :

$$y_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad (\text{II.20})$$

II.3.1-a Estimation du modèle par la méthode des moindres carrés ordinaires (MCO)

Pour analyser la relation entre x et y , la première étape est de tracer un nuage de points (x_i, y_i) pour observer leur tendance. Différents types de relations peuvent exister.

Si les points semblent suivre une relation linéaire (affine), l'objectif est de trouver la droite qui représente au mieux cette tendance. La méthode la plus utilisée pour cela est celle des moindres carrés ordinaires, introduite par Legendre en 1805. Cette méthode consiste à tracer la droite qui minimise la somme des carrés des écarts entre les points observés et cette droite (Figure II.3).

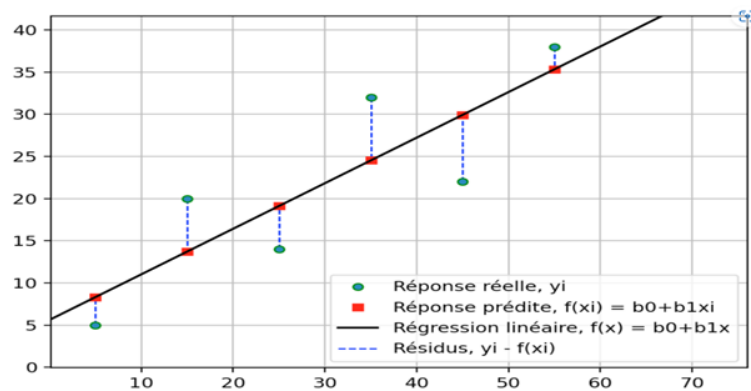


FIGURE II.3 – Exemple de liaisons affine entre x et y [42]

On cherche donc à déterminer les valeurs de $\hat{\beta}_0$ et $\hat{\beta}_1$ qui réduisent au minimum la somme des carrés des erreurs, définie par :

$$S(\hat{\beta}_0, \hat{\beta}_1) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (\text{II.21})$$

ou :

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

Ces valeurs sont appelées estimateurs des moindres carrés ordinaires (MCO).

Expression des estimateurs $(\hat{\beta}_0, \hat{\beta}_1)$:

A la fin des calculs, on obtient pour expressions des estimateurs [38], [40], [41], [44]

$$\hat{\beta}_1 = \sum_{i=1}^n \frac{(x_i - \bar{x})(y_i - \bar{y})}{(x_i - \bar{x})^2} = \frac{(x, y)}{(x)} = \frac{s_{xy}}{s_x} \quad (\text{II.22})$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (\text{II.23})$$

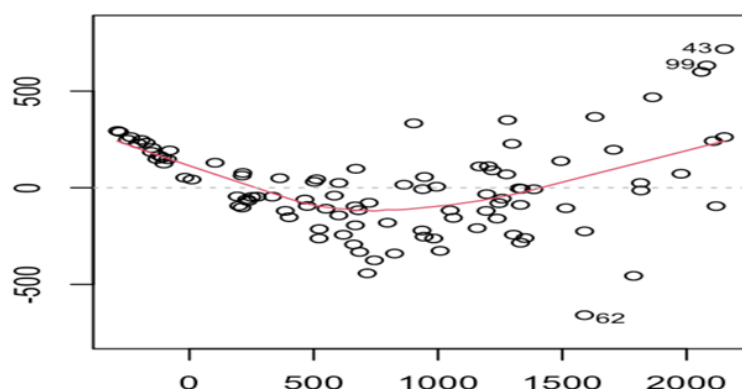


FIGURE II.4 – Exemple de liaisons non affine entre x et y [43]

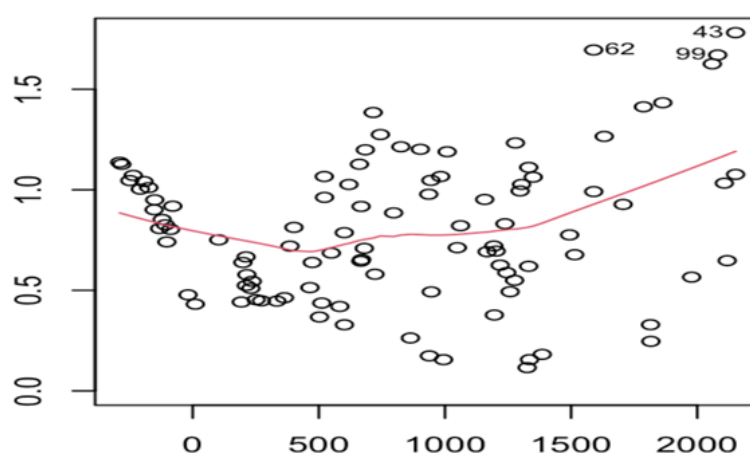


FIGURE II.5 – Exemple pas de liaisons apparente entre x et y [43]

II.3.1-b Test de la pente de la droite de régression

Après avoir postuler notre modèle et estimer ses paramètres, on doit forcément se demander ce que vaut ce modèle en vérifiant son fonctionnement. Autrement dit si cette pente qui représente la liaison entre les x et les y est suffisamment abrupte pour pouvoir affirmer la liaison, il va donc falloir utiliser un test statistique pour vérifier cela. Le test de significativité de la pente consiste à vérifier l'exogène sur l'endogène, qu'on va le voir un peu plus en détail dans la partie RLM.

II.3.2 Régression linéaire multiple

La régression linéaire multiple est une méthode statistique largement employée dans divers domaines pour analyser des données multidimensionnelles et d'estimer la valeur future d'une variable cible (dépendante) en fonction de plusieurs facteurs explicatifs (indépendants). Elle repose sur l'hypothèse qu'il existe une relation linéaire entre ces variables, ce qui permet d'élaborer un modèle mathématique capable de faire des prévisions. Elle représente une extension naturelle de la régression simple.

II.3.2-a Modèle de la RLM

On cherche à décrire la relation existante entre une variable quantitative (Y) appelée variable à expliquer (ou encore, réponse, exogène ou dépendante) et plusieurs variables quantitatives $X_1 \dots X_p$ dites variables explicatives (ou encore de contrôle, endogènes, indépendantes ou régresser).

Le modèle de régression linéaire multiple, noté par RLM, est défini par :

$$Y_i = \beta_o + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip} + \epsilon_i \quad (\text{II.24})$$

Où :

$i=1, \dots, n$: correspond au numéro des observations

y_i : La i^{me} observation de la variable Y : dépendante (puissance d'un panneau).

x_{ij} : est la i -ème observation de la j -ème variable indépendantes (par exemple, rayonnement solaire, température).

β_o : Le coefficient 0 est un paramètre appelé ordonnée à l'origine qui représente la moyenne des lorsque la valeur de chaque variable explicative est égale à 0...

$\beta_1, \beta_2, \dots, \beta_n$: Coefficients de régression ou les coefficients inconnus du modèle que l'on veut estimer à partir des données. (Mesurer l'influence de la variance des X (irradiation, température extra) sur la puissance prédire y)

ϵ_i : est l'erreur du modèle (bruit). Il représente la déviation entre ce que le modèle prédit et la réalité.

II.3.2-b Notation matricielle

Ce modèle s'écrit sous la forme des équations comme suit :

$$\left\{ \begin{array}{l} y_1 = \beta_o + \beta_1 X_{11} + \beta_2 X_{12} + \dots + \beta_p X_{1p} + \epsilon_1 \\ y_2 = \beta_o + \beta_1 X_{21} + \beta_2 X_{22} + \dots + \beta_p X_{2p} + \epsilon_2 \\ \vdots \\ y_n = \beta_o + \beta_1 X_{n1} + \beta_2 X_{n2} + \dots + \beta_p X_{np} + \epsilon_n \end{array} \right\} \quad (\text{II.25})$$

Sous la forme matricielle :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix} \times \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix} \quad (\text{II.26})$$

De façon équivalente, on écrit :

$$Y = \beta X + \epsilon$$

Remarque : Les coefficients $\beta_j (j = 1, \dots, p)$ représentent le changement subi par y_i correspondant à un changement unitaire dans la valeur de la variable explicative, lorsque les autres variables explicatives demeurent inchangées

II.3.2-c Estimation du modèle par la méthode des moindres carrés ordinaires (MCO)

Le vecteur $\beta = (\beta_0, \dots, \beta_p)^t$ et σ_δ^2 sont estimés par minimisation du critère des moindres carrés ordinaires (MCO).

Le principe des moindres carrés choisit le vecteur $\hat{\beta}$ minimisant la fonction de la somme des carrés des résidus, notée par SCE_r .

L'estimateur par moindres carrés ordinaires (MCO) $\hat{\beta}$ de β dans le modèle de régression linéaire multiple est :

$$\hat{\beta} = (X^t X)^{-1} X^t Y \quad (\text{II.27})$$

II.3.2-d Estimation de la variance des résidus

On peut construire l'estimateur sans biais pour qui est :

$$S_{\hat{\epsilon}}^2 = \sigma_{\hat{\epsilon}}^2 = \frac{SCR}{n - p - 1}$$

ou

$$\hat{\epsilon} = Y - X\hat{\beta}$$

La variance résiduelle mérite d'être examinée car elle sert à deux choses :

- Un intermédiaire de calcul (grâce à la variance résiduelle, on peut faire des tests, calculer les intervalles de confiance,)
- Elle constitue également une incertitude qu'on aura sur chaque mesure.

II.3.2-e Décomposition de la variabilité

Lorsqu'on analyse un phénomène, on exprime souvent son origine en pourcentages, attribuant une partie à un facteur et une autre à un autre facteur. Ce même principe s'applique à l'approche de régression.

Lorsqu'une expérience est répétée plusieurs fois, les résultats obtenus varient et ne sont presque jamais identiques. Ces valeurs fluctuent autour de leur moyenne, qui peut être considérée comme un centre de gravité. Cette dispersion des résultats est appelée écart total ou variabilité totale. L'objectif du modèle de régression est d'expliquer une partie de

cette variabilité, tandis qu'une autre partie demeure inexplicée.

Pour évaluer la capacité explicative du modèle, on procède à une décomposition de la variabilité selon la démarche illustrée dans la (Figure II.6).

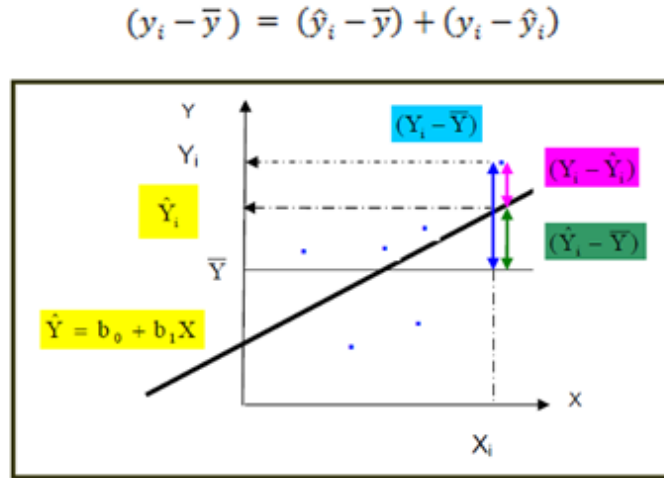


FIGURE II.6 – Explication géométrique de la décomposition de la variabilité.

On peut obtenir la décomposition :

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (\text{II.28})$$

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n \hat{\epsilon}_i^2 \quad (\text{II.29})$$

$$SCE_t = SCE_{reg} + SCE_r \quad (\text{II.30})$$

SCE : désigne la somme des carrés totaux (centrés), autrement dit l'écart total des résultats par rapport à leur moyenne (variabilité totale).

SCE_{reg} : la somme des carrés expliqués (centrés), c'est la portion de variabilité expliquée par le modèle.

SCE_r : la somme des carrés des résidus, c'est la portion inexplicée par le modèle, elle représente l'écart entre ce qu'on observe et ce que le modèle a donné.

II.3.2-f Coefficients de détermination et ajusté

Le rapport entre SCE_{reg} et SCE_t représente la proportion de variance expliquée par le modèle et porte le nom de coefficient de détermination, noté par [26] :

$$R^2 = \frac{SCE_{reg}}{SCE_t} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (\text{II.31})$$

$$R^2 = 1 - \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{SCE_{reg}}{SCE_t} \quad (\text{II.32})$$

Où

SCE_t : désigne la somme des carrés totaux (centrés), autrement dit l'écart total des résultats par rapport à leur moyenne (variabilité totale).

SCE_{reg} : la somme des carrés expliqués (centrés), c'est la portion de variabilité expliquée par le modèle.

SCE_r : la somme des carrés des résidus, c'est la portion inexpliquée par le modèle, elle représente l'écart entre ce qu'on observe et ce que le modèle a donné.

Ce coefficient R^2 est compris entre 0 et 1 : plus il est proche de 1 et plus grande est la part expliquée, autrement dit meilleure est la régression. Inversement, un coefficient R^2 proche de 0 indique que la quantité SCE_r est élevée.

Le coefficient R^2 est un indicateur de la qualité de l'ajustement des valeurs observées par le modèle mais il a le défaut de ne pas tenir compte du nombre de variables explicatives utilisés dans le modèle. On ne peut pas l'utiliser pour comparer plusieurs modèles entre eux car, si on ajoute une variable explicative à un modèle, la part des erreurs diminue forcément et donc le coefficient R^2 augmente : cela signifie que plus il y a de variables explicatives et plus le R^2 est élevé. Or un modèle n'est pas nécessairement meilleur parce qu'il a plus de variables explicatives.

On définit donc un coefficient R^2 ajusté qui tient compte des degrés de liberté. Ce coefficient, noté par R_{adj}^2 , est défini comme suit :

$$R_{adj}^2 = \frac{\frac{SCR}{n-p-1}}{\frac{SCT}{n-1}} = 1 - \frac{n-1}{n-p-1}(1-R^2) \quad (\text{II.33})$$

II.3.2-g Prédiction

Comme dans le cadre du modèle de régression linéaire multiple, les buts de la régression sont de proposer des prédictions pour la variable à expliquer Y lorsque on a de nouvelles valeurs [45] :

$$x_{n+1} = (x_{n+11}; x_{n+1p})^t \quad (\text{II.34})$$

Soit donc :

$$x_{n+1} = (x_{n+11}; x_{n+1p})^t \quad (\text{II.35})$$

Une nouvelle valeur pour laquelle, on veut prédire y_{n+1} qui est définie par :

$$y_{n+1} = x_{n+1}^t \beta + \epsilon_{n+1} \quad (\text{II.36})$$

A partir des n observations précédentes, on a pu calculer un estimateur $\hat{\beta}$ de β . Donc, il est naturel de prédire la valeur correspondante via le modèle ajusté b y_{n+1} définie par :

$$\hat{y}_{n+1} = x_{n+1}^t \hat{\beta} \quad (\text{II.37})$$

II.3.3 Filtre de kalman

Le filtre de kalman est un filtre récursif, qui utilise une approche temporelle [52]. Il permet d'estimer l'état (cet état correspond à l'information utile) d'un système dynamique à partir d'une information, à priori, sur l'évolution de cet état (modèle) et de mesure réelle [53]. L'utilisation de ce filtre est particulièrement recommandée lorsque l'imprécision sur les données traitées peut être interprétée comme un bruit aléatoire, dont la distribution est connue à priori. Le filtre de kalman est un outil de traitement utilisé dans une large gamme de domaines technologiques tels que le traitement du signal, l'automatique le radar la prédiction d'énergie et les systèmes de communication. Il est également de plus en plus utilisé en dehors du domaine du traitement du signal, par exemple en météorologie et en océanographie, en finance et en navigation.

Dans la littérature on trouve plusieurs variantes du filtre de kalman :

- Filtre de kalman standard (KF)
- Filtre de kalman étendu (EKF)
- Filtre de kalman non parfumé (UKF)

On s'intéresse sur le filtre de kalman standard dans ce chapitre

II.3.3-a Historique

Le filtre de kalman introduit par Rudolf Emil kalman en 1960, [54], est l'un des développements mathématiques les plus intéressants dans la théorie de l'estimation linéaire. Il est un reconstituteur d'état dans un environnement stochastique, lorsque les variances des bruits sont connues, c'est un estimateur linéaire minimisant la variance de l'erreur d'estimation. Les applications du filtre de kalman sont nombreuses. Le filtre de kalman permet de donner un estimé de l'état du système à partir d'une information a priori sur l'évolution de cet état (modèle) et des mesures réelles. Il sera utilisé pour estimer des conditions initiales inconnues, prédire des trajectoires, localiser des engins, implanter des lois de commande, etc. [55].

II.3.3-b Présentation du filtre de kalman standard

Le principe d'un tel filtre repose sur un modèle à variables d'état basé sur deux équations [56] :

- L'équation récursive de modélisation du processus aléatoire X_K (vecteur d'état) discret, dite équation d'état ou équation de transition.
- L'équation d'observation (ou équation de mesure) du processus Y_K (vecteur de mesure). Les équations de base du filtre de kalman sont alors les suivantes [57] :

$$\left\{ \begin{array}{l} x_{k+1} = A_k X_k + B_k U_k \\ Y_k = C_k X_k + V_k \end{array} \right\} \quad (\text{II.38})$$

$$W_k \rightarrow N(0, Q_k)$$

$$V_k \rightarrow N(0, R_k)$$

Avec :

X_k : Vecteur d'état à l'instant K qui inclut les quantités à estimer, de taille $n \times 1$.

A_k : Matrice de transition. Elle décrit l'évolution du vecteur d'état de l'instant $K - 1$ à l'instant, de taille $n \times n$.

B_k : Matrice de commande à l'instant K , dépend de la modélisation du système.

C_k : Matrice d'observation (mesure). Elle est en fait le lien entre les paramètres du système et les mesures de taille $m \times n$.

U_k : Vecteur représentant les commandes appliquées au système à l'instant K .

Y_k : Vecteur de mesure à l'instant K , de taille $m \times 1$.

W_k : Bruit de modélisation lié à l'incertitude que l'on sur le modèle de processus.

V_k : Bruit de mesure, de taille $m \times 1$.

Q_k : Matrice de variances-covariances du bruit de processus à l'instant K .

R_k : Matrice de variances-covariances du bruit de mesure à l'instant. Les bruits sont considérés comme blancs, gaussiens, centrés, indépendants entre eux et indépendants de l'état du système.

II.3.3-c Fonctionnement du filtre de kalman

L'estimation optimale consiste à trouver la meilleure estimation \hat{x}_k de l'état x_k en minimisant un critère qui est la variance de l'erreur d'estimation :

$$e_k = x_k - \hat{x}_k$$

$$P_k = E(e_k e_k^T)$$

Tel que e_k : c'est l'erreur et P_k : la covariance de l'erreur.

Filtre de kalman, travaille généralement en deux étapes Prédiction et mise à jour [58]. La phase de prédiction utilise l'état estimé de l'instant précédent pour produire une estimation courante. Dans l'étape de mise à jour, les observations de l'instant courant sont utilisées pour corriger l'état prédit dans le but d'obtenir une estimation plus précise [59].

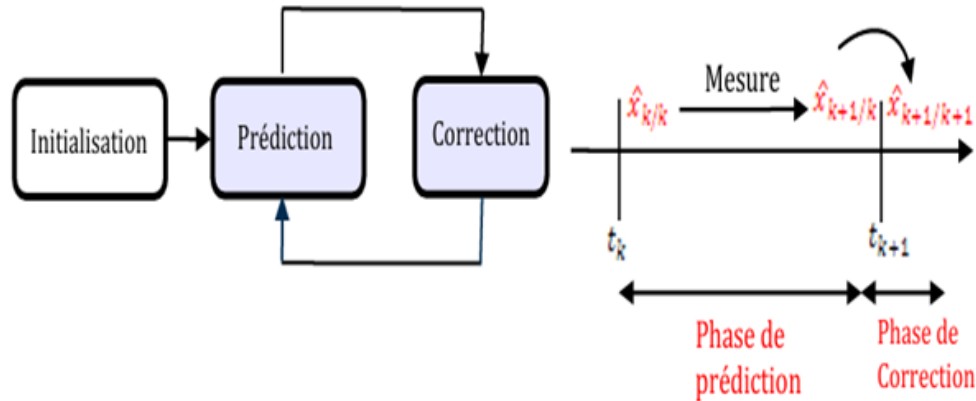


FIGURE II.7 – Etapes du filtre de kalman

1. La phase de prédiction

On se place à l'instant K . A cet instant, on dispose d'une estimation initiale fondée sur la connaissance du processus et des mesures jusqu'à l'instant précédent, c'est-à-dire $K - 1$. Cette estimation est appelée estimation à priori. Si on notera $\hat{x}_{k/k}$ l'estimation d'état à priori, donc l'erreur à priori est donnée par [60] :

$$e_{k/k} = x_k - \hat{x}_{k/k-1} \quad (\text{II.39})$$

Ainsi que la matrice de covariance de l'erreur à priori donnée par :

$$P_{k/k-1} = E(e_{k/k-1}e_{k/k-1}^T) = E((x_k - \hat{x}_{k-1})(x_k - \hat{x}_{k-1})^T) \quad (\text{II.40})$$

Les équations de la phase de prédiction sont :

$$\begin{cases} \hat{x}_{k/k-1} = A_{k-1}X_{k-1/k-1} + B_{k-1}U_{k-1} \\ P_{k/k-1} = A_{k-1}\hat{x}_{k-1/k-1}A_{k-1}^T + Q_{k-1} \end{cases} \quad (\text{II.41})$$

2. La phase de mise à jour (correction)

Nous allons, maintenant, utiliser la mesure Y_k pour corriger l'estimation à priori $\hat{x}_{k/k-1}$ et ainsi obtenir l'estimation à posteriori $\hat{x}_{k/k}$.

L'erreur à posteriori est :

$$e_{k/k} = x_k - \hat{x}_{k/k}$$

La matrice de covariance de l'erreur à posteriori s'écrit alors :

$$P_{k/k} = E(e_{k/k}e_{k/k}^T) = E((x_k - \hat{x}_k)(x_k - \hat{x}_k)^T) \quad (\text{II.42})$$

Les équations de la phase de correction sont :

$$\left\{ \begin{array}{l} \hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k(y_k - c_k \hat{x}_{k/k-1}) \\ P_{k/k} = (I_k - K_k C_k) P_{k/k-1} \end{array} \right\} \quad (II.43)$$

I : est la matrice identité de même taille que $P_{k/k-1}$

K : Gain de Kalman optimal.

II.3.3-d Les Hypothèses

Nous supposons que :

H_1 : La paire (A,c) AC est détectables, c'est-à-dire qu'il n'y a pas de mode instable et inobservable dans le système [61].

H_2 : les signaux $w(t)$ et $v(t)$ sont des bruits blancs gaussiens centrés de Densité Spectrale de Puissance (DSP) W et V respectivement, c'est-à-dire :

$$E(W_k V_i^T) = 0, \forall i, k$$

$$E(W_k W_i^T) = Q_k \delta_{ik}$$

$$E(V_k V_i^T) = R_k \delta_{ik}$$

H_3 : V est inversible (il y a autant de sources de bruits blancs indépendantes que de mesures dans l'équation de mesure).

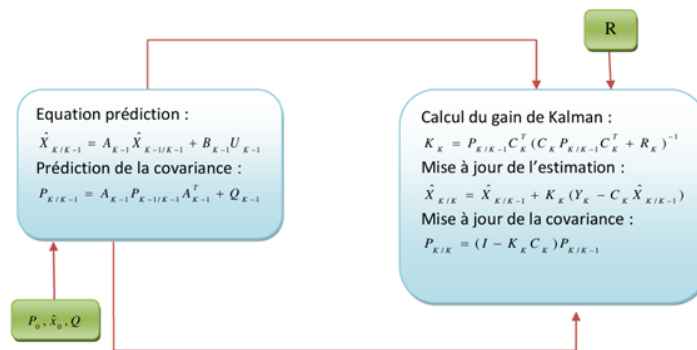


FIGURE II.8 – Processus itératif du filtre de kalman linéaire.

II.3.3-e Algorithme du Filtre de kalman standard

L'algorithme suivant indique les différentes étapes qui permettent d'obtenir un état estimé à chaque instant.

1. Initialisation (conditions initiales) : Lors de l'exécution de l'algorithme il est nécessaire d'initialiser les grandeurs suivantes :

- Le Vecteur d'état initial,
 - La Matrice de covariance d'erreur associée au vecteur d'état initial,
 - La Matrice de covariance des bruits du système,
 - La Matrice de covariance des bruits de mesures,
 - La Matrice de transition d'état,
 - La Matrice de mesure
2. Calcul de l'estimé $\hat{x}_{k/k-1}$ et de la matrice de covariance $P_{k/k-1}$ à partir des mesures disponibles à l'instant $K - 1$

$$\begin{cases} \hat{x}_{k/k-1} = A_{k-1}X_{k-1/k-1} + B_{k-1}U_{k-1} \\ P_{k/k-1} = A_{k-1}\hat{x}_{k-1/k-1}A_{k-1}^T + Q_{k-1} \end{cases} \quad (\text{II.44})$$

3. Calcul du gain de kalman :

$$K_k = P_{k/k-1}C_k^T(C_kP_{k/k-1}C_k^T + R_k)^{-1} \quad (\text{II.45})$$

4. Mise à jour de l'estimation :

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k(Y_k - C_k\hat{x}_{k/k-1}) \quad (\text{II.46})$$

5. Calcul de la matrice de covariance de l'erreur :

$$P_{k/k} = (I_k - K_kC_k)P_{k/k-1} \quad (\text{II.47})$$

6. Préparation du pas suivant [40] :

$$\begin{cases} \hat{x}_{k+1/k} = A_kX_{k/k} + B_kU_k \\ P_{k+1/k} = A_kP_{k/k}A_k^T + Q_k \end{cases} \quad (\text{II.48})$$

II.4 méthodes d'intelligence artificielle

Les méthodes d'intelligence artificielle (IA) visent à reproduire certains mécanismes de l'intelligence humaine afin de résoudre des problèmes complexes, souvent non linéaires, difficiles à modéliser par des approches classiques.

Dans le domaine de la prévision de la puissance photovoltaïque, ces méthodes permettent d'apprendre à partir de grandes quantités de données, d'adapter les modèles aux variations des conditions, et d'améliorer les performances prédictives. Parmi ces techniques, on retrouve notamment les réseaux de neurones artificiels, les systèmes flous, et les algorithmes d'apprentissage automatique. Dans cette partie on s'intéresse sur réseaux de neurones artificielle

II.4.1 Réseaux de neurones

Les réseaux de neurones artificiels (RNA) sont des modèles qui imitent le fonctionnement du cerveau humain. Ils s'inspirent de la manière dont les neurones traitent l'information. L'idée principale est que l'intelligence humaine vient de la structure du système nerveux. En reproduisant ce fonctionnement sur un ordinateur, on peut le rendre plus "intelligent", tout en gardant une bonne puissance et une exécution rapide [46].

II.4.1-a Historique

Le développement des réseaux de neurones artificiels a suivi une évolution marquée par plusieurs étapes clés, depuis les premières modélisations théoriques jusqu'aux applications modernes de l'intelligence artificielle.

— **1943 : Le neurone formel (McCulloch & Pitts)**

McCulloch et Pitts proposent le premier modèle mathématique de neurone artificiel. Ce neurone formel est capable d'effectuer des opérations logiques simples comme les fonctions logiques ET, OU et NON. Il s'agit de la première abstraction mathématique inspirée du fonctionnement du neurone biologique.

— **1949 : Première règle d'apprentissage (Hebb)**

Donald Hebb introduit une règle d'apprentissage basée sur une hypothèse biologique : "Les neurones qui s'activent ensemble se renforcent ensemble". Cette idée servira de base aux mécanismes d'apprentissage synaptique dans les réseaux de neurones.

— **1958 : Le perceptron (Rosenblatt)**

Frank Rosenblatt développe le perceptron, un modèle de neurone artificiel capable d'apprendre à classer des données en ajustant ses poids synaptiques. C'est un réseau simple à une seule couche, capable de résoudre des problèmes linéairement séparables.

— **1960 : L'adaline (Widrow & Hoff)**

Widrow et Hoff proposent l'Adaline (Adaptive Linear Neuron), un modèle proche du perceptron, mais utilisant une règle d'apprentissage fondée sur la descente du gradient. Ce modèle introduit le concept d'erreur de prédiction pour ajuster les poids, ouvrant la voie à des méthodes d'optimisation plus efficaces.

— **1969 : Limites du perceptron (Minsky & Papert)**

Marvin Minsky et Seymour Papert publient un ouvrage démontrant les limitations du perceptron, notamment son incapacité à résoudre des problèmes non linéaires tels que la fonction XOR. Cette découverte met en lumière la nécessité de réseaux à plusieurs couches. Cependant, à cette époque, il n'existe pas de méthode efficace pour entraîner ces architectures plus complexes. Cette impasse technique provoque un ralentissement majeur des recherches dans ce domaine.

— **~ 1969 – 1986 : Période de stagnation**

Suite aux critiques de Minsky et Papert, la recherche sur les réseaux de neurones connaît une période de stagnation d'environ 17 ans. Faute de solutions concrètes pour entraîner les réseaux multicouches, les travaux dans ce domaine ralentissent considérablement.

— **1986 : Rétropropagation (Rumelhart & McClelland)**

La publication de l'algorithme de rétropropagation du gradient par Rumelhart et McClelland marque une avancée décisive. Cet algorithme permet enfin d'entraîner des réseaux multicouches, en ajustant les poids de manière efficace en fonction de l'erreur de sortie. Cela relance fortement l'intérêt pour les réseaux de neurones.

— **Actuellement : Multiples applications**

Aujourd'hui, grâce aux progrès informatiques (augmentation de la puissance de calcul, disponibilité de grandes quantités de données), les réseaux de neurones, notamment profonds (deep learning), connaissent un développement spectaculaire. Ils sont utilisés dans des domaines variés : reconnaissance d'images, traitement automatique du langage, diagnostic médical, véhicules autonomes, etc.

II.4.1-b Réseau de neurones : fondement biologique

C'est l'élément de base du système nerveux. Il reçoit des signaux en provenance des neurones voisins d'un côté par des chevelures appelées dendrites, les traite dans le corps cellulaire puis engendre, conduit et transmet l'influx nerveux de l'autre côté à d'autres neurones via un long prolongement appelé axone. La communication entre deux neurones voisins s'effectue au niveau de la terminaison de l'axone, cet endroit est appelé synapse.

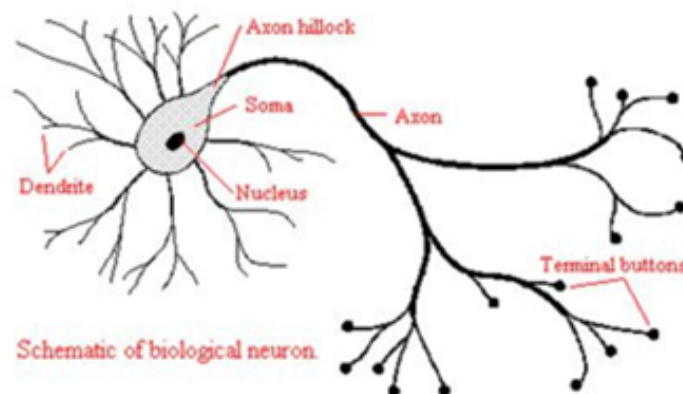


FIGURE II.9 – Neurone biologique [47]

II.4.1-c Réseaux de neurones artificiels

1. Principe général

Le neurone artificiel est inspiré du neurone biologique :

- Dans le neurone biologique, l'information passe par les synapses, traverse le corps cellulaire et est transmise via l'axone.

- Dans le neurone artificiel, les entrées sont pondérées, une fonction de transfert traite les données, puis une sortie est produite. Les deux types de neurones fonctionnent de manière similaire en transmettant des informations d'un point à un autre.

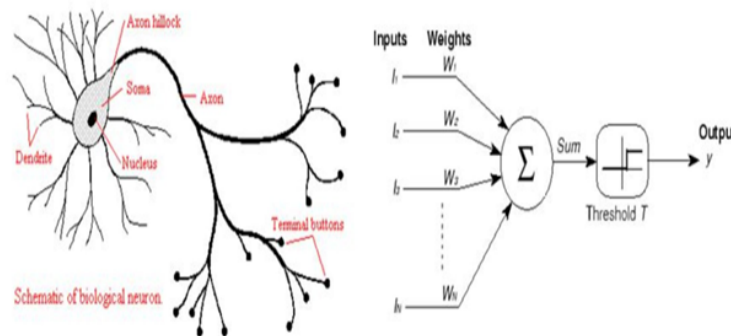


FIGURE II.10 – Neurone biologique / neurone artificiel. [47]

2. Modèle mathématique du neurone artificiel (formel)

Un neurone formel simple peut réaliser des fonctions logiques, arithmétiques et symboliques complexes. Alors qu'est-ce qu'un neurone formel ?

Un neurone artificiel formel est, comme on l'a vu, l'unité élémentaire de traitement d'un réseau de neurones. Il est connecté à des sources d'information en entrée (d'autres neurones par exemple) et renvoie une information en sortie.

— Entrées

Un neurone artificiel reçoit plusieurs informations x_i avec $1 \leq i < n$. Chaque information est multipliée par un poids w_i , qui indique son importance.

Il existe aussi un biais b . L'ensemble des informations est combiné sous forme de somme pondérée :

$$s = \sum_{i=0}^n w_i x_i = \left(\sum_{i=1}^n w_i x_i \right) - b \quad (\text{II.49})$$

Cette valeur S est ensuite envoyée à une fonction d'activation.

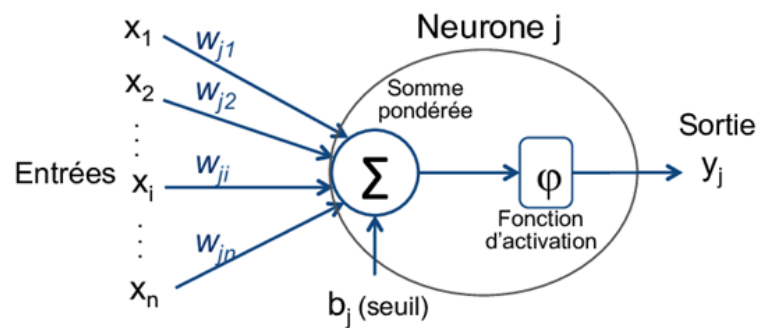


FIGURE II.11 – Neurone Artificiel [48]

— **Fonction d'activation**

Cette donnée est passée à une fonction f dite (fonction d'activation) qui représente un filtre permettant d'adapter la valeur de la somme précédente aux caractéristiques de la sortie désirée. C'est une fonction qui, généralement, doit renvoyer un réel proche de 1 quand les "bonnes" informations d'entrée sont données et un réel proche de 0 quand elles sont "mauvaises". La valeur de la fonction d'activation est la sortie y du neurone [49].

$$y = f(s) = f\left(\left(\sum_{i=1}^n w_i x_i\right) - b\right) \quad (\text{II.50})$$

Où :

b : le biais

w_i : le poids synoptiques

x_i : les entres

Il existe beaucoup de fonctions d'activations possibles, toutefois dans la pratique il y en a principalement deux qui sont utilisées [49].

— La fonction de Heaviside est définie par :

$$\forall x \in \mathbb{R} \quad f(x) = 1 \quad \text{si } x \geq 0, 0 \text{ sinon}$$

— La fonction de sigmoïde est définie par :

$$\forall x \in \mathbb{R} \quad f(x) = \frac{1}{1 + e^{-x}}$$

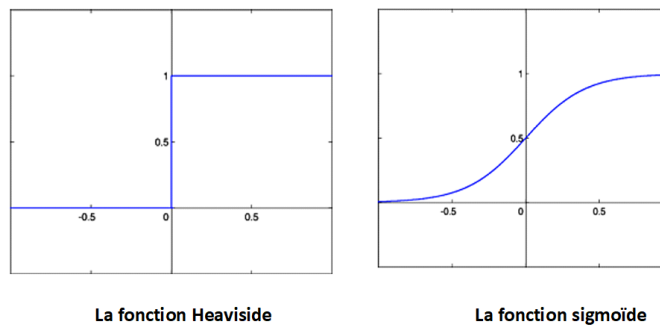


FIGURE II.12 – Fonctions de transfert fréquemment utilisées

II.4.1-d Différents modèles de RNA

Les réseaux de neurones artificiels (RNA) se divisent en deux grandes catégories.

1. **Réseaux de neurones directs** Dans ce type de réseau (Figure II.13), l'information circule toujours dans un seul sens. Chaque neurone d'une couche $n1$ est connecté à tous les neurones de la couche suivante n .

Il existe plusieurs sous-types :

- Perceptron monocouche : une seule couche cachée.
- Perceptron multicouche : plusieurs couches cachées.
- Neurone à base radiale : utilise une fonction de transfert de type gaussienne.

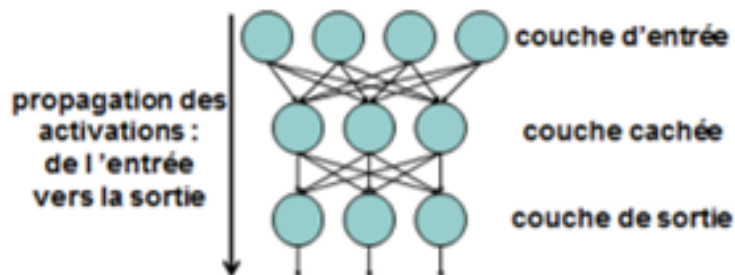


FIGURE II.13 – IRéseaux directs (feedforward networks)

2. Réseaux de neurones récurrents (feedback networks)

Dans ces réseaux, la sortie d'un neurone d'une couche n peut servir d'entrée pour un neurone d'une couche précédente $n1$ ou un neurone de la même couche n [50].

On distingue plusieurs sous-types :

- Réseaux compétitifs
- Cartes de Kohonen
- Réseaux de Hop Field
- Modèles ART

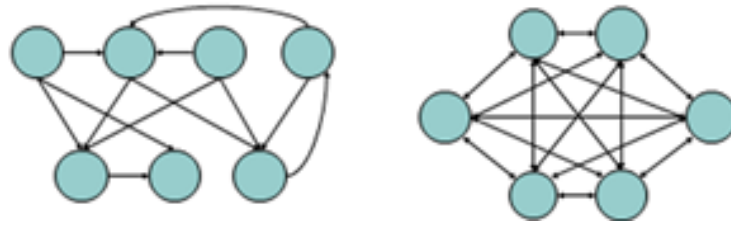


FIGURE II.14 – Réseaux récurrents (feedback networks).

II.4.1-e Méthodes d'apprentissage ou d'entraînement

Comme le cerveau humain, les réseaux de neurones artificiels (RNA) peuvent apprendre par expérience. L'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré. Trois grandes classes d'apprentissage existent :

1. Apprentissage supervisé (back propagation)

Cet algorithme d'apprentissage (Figure II.15) est applicable uniquement lorsque les associations entre les entrées et les sorties désirées sont définies à l'avance. Son utilisation facilite l'apprentissage et le rend considérablement plus rapide que les deux autres méthodes, car l'ajustement des poids s'effectue directement en fonction de l'erreur, c'est-à-dire la différence entre la sortie obtenue par le réseau de neurones artificiels (RNA) et la sortie attendue [50].

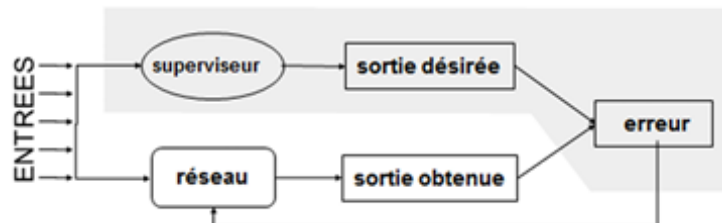


FIGURE II.15 – Processus d'apprentissage supervisé.

2. Apprentissage non-supervisé

Dans ce cas, les sorties désirées ne sont pas connues à l'avance pour les entrées données (Figure II.16). L'apprentissage repose sur une approche exploratoire où l'algorithme ajuste dynamiquement les poids des connexions entre les neurones afin d'optimiser la qualité de la classification des entrées [50].

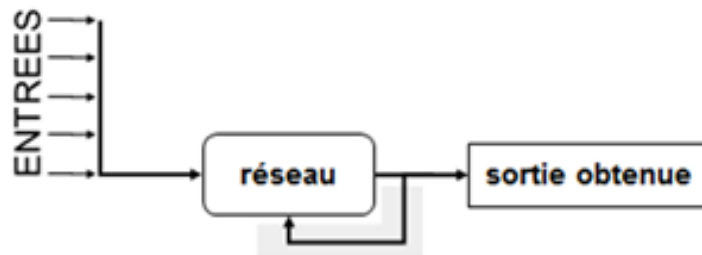


FIGURE II.16 – Processus d'apprentissage non-supervisé.

II.4.1-f Réseau de neurones monocouche

Un perceptron est un neurone artificiel doté d'une règle d'apprentissage permettant d'ajuster automatiquement les poids synaptiques afin de résoudre un problème d'apprentissage supervisé.

Le perceptron est composé d'un nombre variable de neurones disposés verticalement sur une seule couche, formant ainsi un réseau de neurones monocouche (Figure 17).

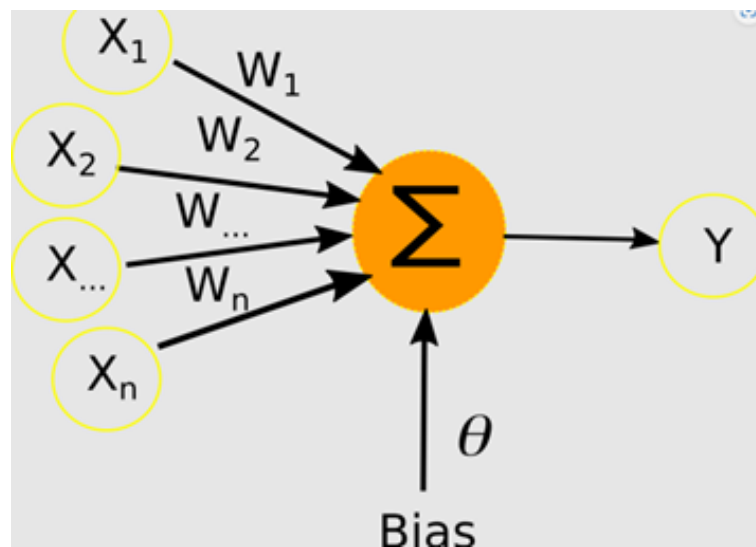


FIGURE II.17 – structure de Perceptron de Rosenblatt (1958) [51]

II.4.1-g Lois d'apprentissage

Lors de la phase d'apprentissage d'un réseau de neurones, l'objectif est d'optimiser les performances du réseau de neurones artificiels (RNA) afin qu'il puisse résoudre un plus grand nombre de tâches.

C'est dans cette optique que les algorithmes d'apprentissage interviennent en ajustant les poids w_{ij} reliant les neurones ainsi que les biais.

De manière générale, la mise à jour d'un poids de connexion s'effectue selon la règle suivante :

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n) \quad (\text{II.51})$$

1. Loi de Hebb :

Cette règle d'apprentissage supervisé, relativement simple, repose sur l'hypothèse que lorsqu'un neurone " *i* " est activé de manière répétitive ou persistante par un neurone " *j* ", alors le poids synaptique associé $w_{ij}(n)$ est modifié. Cette modification peut être une augmentation ou une diminution d'une quantité $\Delta w_{ij}(n)$, qui correspond à une constante positive représentant l'intensité de l'apprentissage (Learning rate).

$$w_{ij}(k+1) = w_{ij}(k) + \Delta w_{ij}(k) \quad (\text{II.52})$$

$$\Delta w_i(k) = \mu y_d(k) x_i(k) \quad (\text{II.53})$$

Avec :

μ : le taux d'apprentissage,

$y_d(k)$: La réponse désirée ou attendue par le réseau,

$x_i(k)$: l'entrée du neurone *i*.

2. Loi de Rosenblatt ou du perceptron

La règle de Hebb peut conduire à des divergences dans certains cas, bien qu'une solution existe.

Pour pallier cette limitation, une autre règle d'apprentissage a été introduite, prenant en compte l'erreur observée $e(k)$ en sortie. Il s'agit de la règle du perceptron, qui s'exprime ainsi :

$$w_i(k+1) = w_i(k) + \Delta w_i(k) \quad (\text{II.54})$$

où

$$\Delta w_i(k) = \mu [y_{dsire}(k) - y_{calcule}(k)] x_i(k) \quad (\text{II.55})$$

$$\Delta w_i(k) = \mu e(k) x_i(k) \quad (\text{II.56})$$

3. Règle de Widrow-Hoff (1960) (delta-rule ou méthode des moindres carrées (LMS : Least Mean Square))

La règle de Widrow-Hoff constitue une version adaptée de la loi de Hebb. Elle repose sur une mise à jour continue des poids synaptiques entre les neurones, dans le but de minimiser l'écart (appelé delta) entre la sortie attendue et la sortie produite par

le neurone.

Cette règle s'exprime par la relation suivante :

$$w_i(k+1) = w_i(k) + \mu[d(k) - y(k)]x_i(k) \quad (\text{II.57})$$

Où μ est le taux d'apprentissage, $d(k)$ la sortie désirée, $y(k)$ la sortie obtenue, et $x_i(k)$ l'entrée du neurone i .

La règle de Widrow-Hoff présente ainsi une grande similarité avec la règle d'apprentissage du perceptron.

II.4.1-h Réseau de neurones multicouche : perceptron multicouches (Multi-Layers Perceptron (MLP))

1. Architecture d'un réseau MLP

Comme nous l'avons déjà mentionné, le perceptron à une seule couche ne peut apprendre efficacement que lorsque les classes à distinguer sont linéairement séparables. Pour dépasser cette limitation, l'idée a émergé d'utiliser un réseau de type MLP (Multilayer Perceptron).

Le perceptron multicouche représente aujourd'hui l'une des architectures de réseaux de neurones les plus répandues. Il se compose d'au moins trois couches de neurones. Dans cette structure, les neurones d'une couche ne sont connectés qu'aux neurones de la couche suivante (voir Figure II.18), ce qui permet à l'activation de se propager progressivement des neurones d'entrée jusqu'à ceux de sortie. [50]

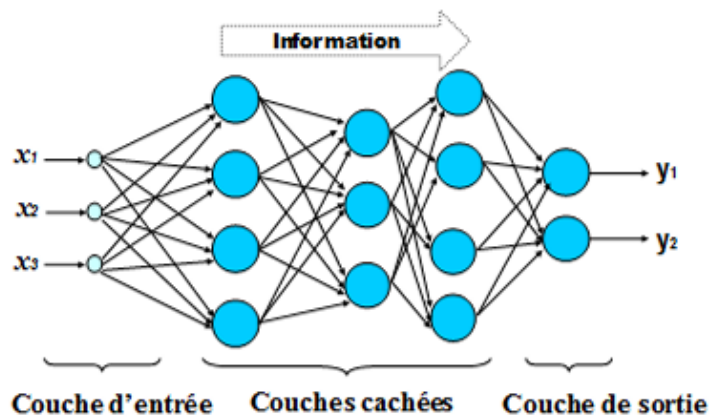


FIGURE II.18 – Perceptron multicouches

2. Processus d'apprentissage dans les réseaux MLP

(a) Algorithme d'apprentissage

Un réseau de neurones artificiel (RNA) développe ses capacités et améliore ses performances en ajustant ses paramètres internes, tels que les poids et les biais, à l'aide d'un algorithme d'apprentissage adapté.

(b) **Algorithme de rétropropagation du gradient (Backpropagation)**

L'algorithme de rétropropagation du gradient est une méthode d'apprentissage supervisé. Il repose sur l'utilisation d'un ensemble de paires entrée-sortie pour ajuster les paramètres du réseau (voir Figure II.19). [50]

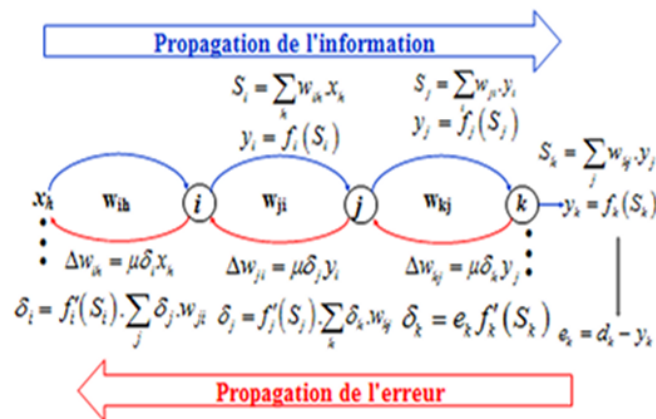


FIGURE II.19 – Schéma explicatif de l'algorithme de rétropropagation du gradient

L'algorithme RPG consiste à mesurer l'erreur entre les sorties désirées $d(n)$ et les sorties Observées $y(n)$.

Il nécessite des fonctions d'activations dérivables des différentes couches de neurones.

L'algorithme RPG consiste à ajuster les poids et biais liés à chaque neurone par la rétropropagation de l'erreur calculée sur le neurone de sortie, tout en commençant par la dernière couche (couche de sortie) vers les couches inférieures jusqu'à arriver à la première (couche d'entrée) Alors ce processus d'apprentissage supervisé se diffère selon l'emplacement du neurone dans le réseau. On peut donc citer deux cas distincts :

— Evaluation du gradient pour un neurone dans la couche de sortie.

Considérons un neurone de sortie k :

L'ajustement des poids se fait toujours à l'aide de l'équation :

$$w_{kj}(n+1) = w_{kj}(n) + \Delta w_{kj}(n) \quad (\text{II.58})$$

La correction à appliquer peut s'écrire sous la forme :

$$\Delta w_{kj}(n) = \mu \delta_k(n) y_j(n) \quad (\text{II.59})$$

où $\delta_j(n)$ est le gradient local défini par :

$$\delta_k(n) = e_k(n)f'(S_k(n)) \quad (\text{II.60})$$

Avec :

$y_j(n)$: sortie calculée par le neurone .

μ : le taux d'apprentissage

$e_k = d_k - y_k$: erreur de sortie

$f'(S_k(n))$: dérivée de la fonction d'activation sur la sortie du neurone.

- Evaluation du gradient pour un neurone dans la couche cachée.
Lorsqu'un neurone j se trouve dans une couche cachée du réseau, il ne dispose pas de sortie désirée directe. Par conséquent, la formule d'ajustement des poids s'exprime différemment, comme suit :

$$\Delta w_{ij}(n) = \mu \delta_j(n) y_i(n) \quad (\text{II.61})$$

Où :

$$\delta_k(n) = f'_j(S_j(n)) \sum_k \delta_k(n) w_{kj}(n)$$

Il est important de noter que, pour la première couche cachée du réseau, en l'absence de couche de neurones en amont, la variable $y_i(n)$ doit être remplacée par l'entrée $x_i(n)$.

- i. Pour une unité k de la couche de sortie O :

$$\delta_k^o(n) = f'_k(S_k(n)) e_k(n) = f'_k(S_k(n)) (d_k(n) - y_k(n)) \quad (\text{II.62})$$

- ii. Pour une unité i de la couche cachée :

$$\delta_i^{-1}(n) = f'_i(S_i(n)) \sum_k \delta_j^e(n) w_{ji}(n) \quad (\text{II.63})$$

3. Optimisation de l'algorithme

Il existe en effet plusieurs stratégies, que ce soit dans le choix des paramètres ou des procédures, permettant d'améliorer de manière significative l'efficacité de l'algorithme de rétropropagation. [50]

- (a) **Modes d'entraînement** : séquentiel et par lot ("batch")

- **Mode séquentiel** (aussi appelé en ligne ou stochastique) :

Dans ce mode, les poids sont mis à jour après chaque exemple d'apprentissage.

- i. Plus rapide, notamment lorsque l'ensemble d'entraînement est volumineux et contient beaucoup de redondances.
- ii. Peut entraîner des difficultés de convergence.

— **Mode par lot ("batch")**

Dans ce cas, les poids sont actualisés uniquement après la présentation de l'ensemble complet des exemples d'apprentissage.

- i. **Avantage** : stabilité dans la mise à jour des poids.
- ii. **Inconvénient** : nécessite plus de calculs et de mémoire lorsque le nombre d'exemples est élevé.

(b) **Correction des oscillations**

La formule de correction des poids :

$$w_{ji}(n+1) = w_{ji}(n) + \mu\delta_j(n)y_i(n) \quad (\text{II.64})$$

Elle ne prend en compte que l'erreur commise à l'instant n , sans considérer les ajustements effectués lors des itérations précédentes.

Un tel apprentissage peut entraîner des variations instables des poids, parfois sous forme d'oscillations, empêchant ainsi la convergence du réseau [50].

Pour remédier à ce problème, on ajoute un terme proportionnel à la précédente modification des poids, ce qui permet de lisser et de stabiliser l'apprentissage [50] ;

$$w_{ji}(n+1) = w_{ji}(n) + \alpha\Delta w_{ji}(n-1) + \mu\delta_j(n)y_i(n) \quad (\text{II.65})$$

Où α , compris entre 0 et 1, est le moment "momentum"

La relation de correction des poids :

$$\Delta w_{ji}(n) = \Delta w_{ij}(n-1) + \mu\delta_j(n)y_i(n) \quad (\text{II.66})$$

Est appelé la "règle delta généralisée" (generalized delta rule).

Pour $\alpha = 0$, on obtient la "règle delta".

II.5 Conclusion

Dans ce chapitre, nous avons présenté les différentes approches statistiques et physiques utilisées pour la prédiction de la puissance photovoltaïque. Ces méthodes permettent d'estimer la production énergétique d'un système photovoltaïque en tenant compte de plusieurs variables climatiques telles que l'irradiation solaire, la température et parfois l'humidité.

SOMMAIRE

III.1 INTRODUCTION	48
III.2 DESCRIPTION TECHNIQUE DU SITE ÉTUDIÉ : CENTRALE PHOTOVOLTAÏQUE D'ADRAR	48
III.3 ANALYSE DES DONNÉES	49
III.3.1 Interprétation des figures	50
III.4 PRÉDICTION DE LA PUISSANCE	52
III.4.1 Application de la méthode RLM	54
III.4.2 Application de réseau de neurone	57
III.4.3 Régression Linéaire Multiple (RLM) combinée avec Réseau Elman (NAR)	61
III.4.4 Régression Linéaire Multiple (RLM) combinée avec Filtre de Kalman . . .	64
III.4.5 Comparaison entre les 4 méthodes	68
III.5 CONCLUSION	70

III.1 Introduction

Dans ce chapitre, nous présentons et analysons les résultats obtenus à travers différentes approches de modélisation appliquées à la prédiction de la puissance photovoltaïque sur le site d'Adrar. Afin d'évaluer l'efficacité des méthodes proposées, nous avons utilisé plusieurs techniques de modélisation, allant de la régression linéaire multiple (RLM) à des approches plus avancées, les réseaux de neurones récurrents (RNN) et les filtres de Kalman.

Nous débutons par l'application de la régression linéaire multiple seule, méthode statistique classique permettant de modéliser la relation linéaire entre la puissance solaire et un ensemble de variables explicatives. Par la suite, nous mettons en œuvre un modèle RNN, reconnu pour sa capacité à capturer les dépendances temporelles dans les séries chronologiques. Nous proposons également deux combinaisons : RLM + RNN, RLM+ filtre de Kalman.

Les résultats de chaque méthode sont présentés sous forme de métriques d'évaluation appropriées et de visualisations graphiques permettant une interprétation claire des performances. Enfin, une comparaison globale entre les trois approches principales est réalisée afin de déterminer la méthode la plus performante et la plus adaptée à la prédiction de la puissance photovoltaïque dans la région étudiée.

III.2 Description technique du site étudié : Centrale photovoltaïque d'Adrar

Le site étudié dans ce travail est la centrale solaire photovoltaïque d'Adrar, située à environ 10 km au sud du centre-ville d'Adrar, dans le sud-ouest algérien (27.9077° N, -0.3174° E), au cœur du désert du Sahara. Cette centrale, exploitée par SKTM (Société de Production d'électricité de Sources Renouvelables), filiale du groupe Sonelgaz, a été mise en service en octobre 2015 dans le cadre du programme national de développement des énergies renouvelables [62, 63].

La centrale couvre une superficie totale d'environ 40 hectares est équipée de 75 000 modules photovoltaïques en silicium polycristallin, chacun ayant une puissance de 265 WC, soit une puissance totale installée de 19,875 MWC [64]. Les modules sont fixés sur des structures métalliques fixes, orientées plein sud avec une inclinaison de 25°, une configuration optimisée pour les conditions de fort ensoleillement de la région saharienne [65].

Le système de conversion comprend 24 onduleurs centralisés d'une puissance nominale de 800 kW chacun, assurant la transformation du courant continu (DC) généré par les modules en courant alternatif (AC) injecté dans le réseau électrique national [66].

Pour assurer le suivi et l'analyse des performances, la centrale est équipée d'un système complet de supervision SCADA (Supervisory Control and Data Acquisition). Celui-ci permet l'acquisition automatique de données techniques et environnementales, incluant :

- L'irradiance solaire (en W/m^2),

- La température ambiante ($^{\circ}\text{C}$),
- La température des modules photovoltaïques ($^{\circ}\text{C}$),
- L'humidité (HR)
- La puissance électrique générée (W),
- Les états de fonctionnement des onduleurs et la qualité de l'énergie injectée au réseau [67].



FIGURE III.1 – parc des panneaux solaire au cite d'Adrar

III.3 Analyse des données

Dans cette partie on va mesurer l'influence de chaque paramètre (température ambiante, température de module, irradiation et l'humidité) sur la puissance.

pour analyser la corrélation entre la puissance est les paramétré précédent on a utilisé la méthode de régression linéaire multiple. Sur une période de 1440 donné de la station adrar. La mesure effectuer pour chaque 15min.

TABLE III.1 – Tableau de donné pour tester l'influence des paramètres sur la puissance

Nombre de données	Température ambiante	Température de module	Humidité	Irradiation	Puissance
1	9.9	105.9	25.6	0	0
2
.
.
38	6	87	38.9	353.4	1768
39	6.6	136.8	38.1	428.5	2179
.
1000	13.2	125.5	35.4	177.7	883
.
1440	19.9	124	21.9	0	0

III.3.1 Interprétation des figures

Afin de mieux comprendre les facteurs influençant la puissance de sortie d'un système photovoltaïque (PV), une régression linéaire multiple (RLM) a été effectuée. Les graphiques suivants illustrent la relation linéaire entre la puissance PV et les variables explicatives telles que la température ambiante, la température de module, l'irradiance, l'humidité, et la puissance mesurée.

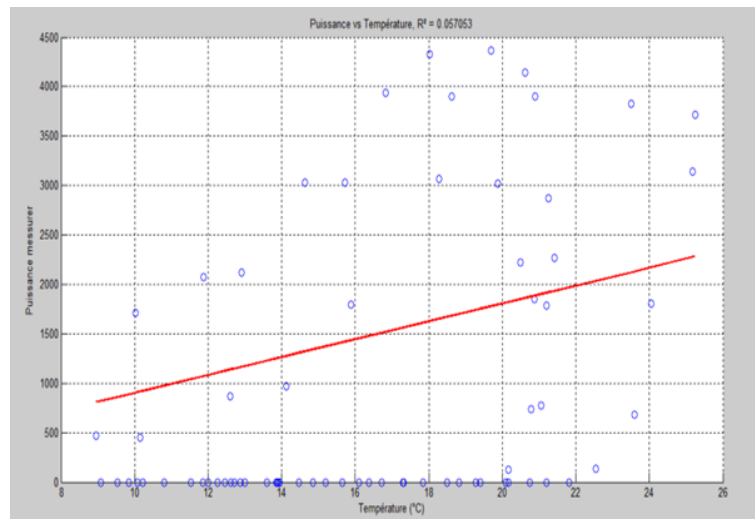


FIGURE III.2 – corrélation entre la température ambiante et la puissance

La (Figure III.2) Montre la relation entre la puissance PV et la température ambiante. Le coefficient de détermination linéaire $R^2=0.0571$ est très faible, ce qui indique que la température a une influence limitée sur la puissance de sortie du système PV. La dispersion importante des points autour de la droite de régression confirme cette faible corrélation. Cela est cohérent avec les connaissances théoriques, selon lesquelles une hausse de température peut même réduire légèrement l'efficacité des panneaux photovoltaïques.

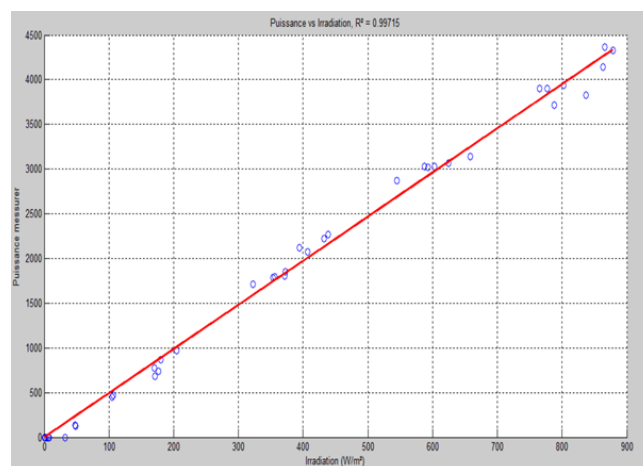


FIGURE III.3 – Corrélation entre l'irradiance et la puissance

La (Figure III.3) représente la relation entre la puissance PV et l'irradiance. Le coefficient de détermination $R^2=0.9971$ est élevé, indiquant une forte corrélation positive. Cela signifie que l'irradiance est un facteur déterminant majeur dans la génération de puissance PV, ce qui est logique, car les panneaux solaires produisent plus d'énergie sous une forte irradiation solaire.

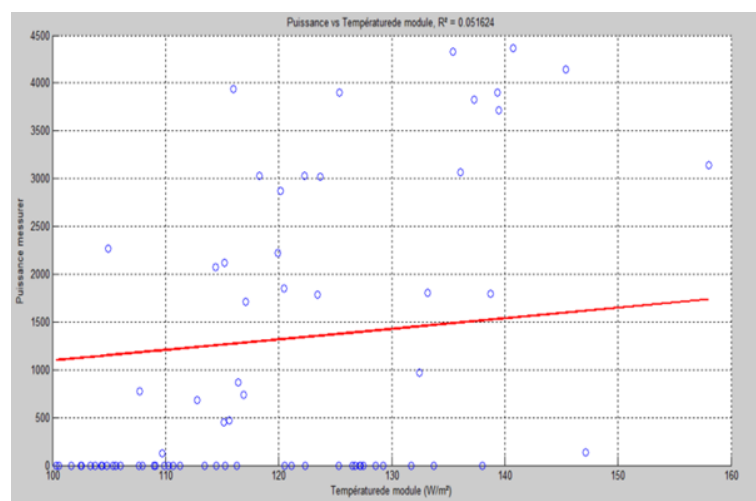


FIGURE III.4 – corrélation entre la température de module et la puissance

La (Figure III.4) Montre la relation entre la puissance PV et la température de module. Le coefficient de détermination linéaire $R^2=0.0515$ est très faible, ce qui indique que la température de module a une influence limitée sur la puissance de sortie du système PV. La dispersion importante des points autour de la droite de régression confirme cette faible corrélation. Cela est cohérent avec les connaissances théoriques, selon lesquelles une hausse de température peut même réduire légèrement l'efficacité des panneaux photovoltaïques.

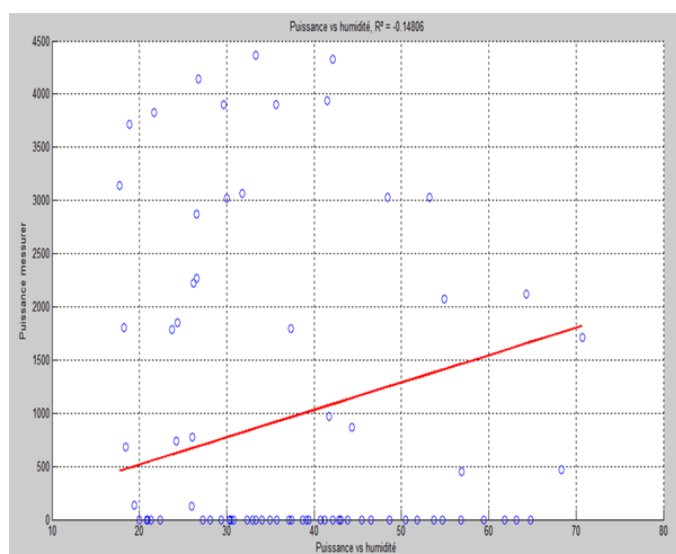


FIGURE III.5 – Corrélation entre l'humidité et la puissance

La (Figure III.5) montre la relation entre la puissance PV et l'humidité. Le coefficient affiché est $R^2=0.1481$. Cette valeur négative, très faible, indique l'absence de relation linéaire significative entre l'humidité et la puissance produite. Ainsi, on peut conclure que l'humidité n'a pratiquement aucune influence directe sur la puissance du système PV dans les conditions observées.

Après cette étude on va annuler l'humidité dans tout le travail suivant.

III.4 Prédiction de la puissance

Dans le cadre de cette étude, un jeu de données complet a été fourni par le système de supervision SCADA de la centrale solaire photovoltaïque d'Adrar. Les données recueillies couvrent une période significative de 105216 données, ce qui permet une observation fine de l'évolution des paramètres environnementaux et électriques influençant la production photovoltaïque.

Les données ont été collectées à une fréquence de 15 minutes, soit 4 mesures par heure. Pour la modélisation, une moyenne horaire a été calculée pour obtenir une seule valeur par heure. Les données principales utilisées sont : température ambiante, irradiation et puissance mesurée.

TABLE III.2 – Tableau de données totale

Nombre de données	Température ambiante	Irradiation	Puissance
1	7,5	0	0
2	7	0	0
.	.	.	.
36	3	138,8	2855
37	3,4	366,1	5321
.	.	.	.
2058	14,8	518,1	8807
.	.	.	.
.	.	.	.
.	.	.	.
105216	14	0	0

Afin de faciliter l'analyse, la modélisation et la visualisation, une agrégation horaire des données a été réalisée. Pour cela, la moyenne de chaque variable a été calculée sur des intervalles horaires en regroupant chaque série de 4 valeurs consécutives ($4 \times 15min = 1h$). Cette opération a permis d'obtenir un nouveau jeu de données horaire comprenant 26 304 valeurs pour chaque paramètre, correspondant à la moyenne horaire de la température ambiante, de la température des modules, de l'irradiance et de la puissance générée.

La valeur moyenne horaire de la puissance $P_m(i)$ produite par la centrale PV :

$$P_m(i) = \frac{1}{N} \sum P(t) \quad (\text{III.1})$$

Avec :

$P(t)$: la puissance instantanée mesurée à chaque 15min

$N = 4$: Nombre de mesure à chaque houer

I : le nombre d'heures ($i=1 : 26304$)

La valeur moyenne horaire de l'irradiation $I_m(i)$ mesurée dans la centrale PV :

$$I_m(i) = \frac{1}{N} \sum I_m(t) \quad (\text{III.2})$$

La valeur moyenne horaire de la température de module $T_m(i)$ mesurée dans la centrale PV :

$$T_m(i) = \frac{1}{N} \sum T_m(t) \quad (\text{III.3})$$

La valeur moyenne horaire de la température ambiante $T_a(i)$ mesurée dans la centrale PV :

$$T_a(i) = \frac{1}{N} \sum T_a(t) \quad (\text{III.4})$$

Le tableau des données horaires est :

TABLE III.3 – Tableau des données horaires d'entraînement

Nombre d'heure	Puissance moyenne	Irradiation moyenne	Température ambiante moyenne
1	0	0	6.6750
2	0	0	6.1250
.	.	.	.
40	13026	795.70	17.6250
.	.	.	.
.	.	.	.
.	.	.	.
20000	730	78.750	19
.	.	.	.
.	.	.	.
26304	0	0	14

On prendra 80 % des données (21043) pour l'entraînement et 20 % (5261) pour le test.

TABLE III.4 – Tableau des données horaires de teste

Nombre d'heures	Puissance moyenne	Irradiation moyenne	Température ambiante moyenne
21043	0	0	5.6750
21044	0	0	6.0250
.	.	.	.
31043	12026	775.30	16.6450
.	.	.	.
.	.	.	.
	610	60.75 0	17
.	.	.	.
105216	0	0	14

III.4.1 Application de la méthode RLM

Formulation mathématique du modèle

$$P_{prdit} = \beta_1 T_a + \beta_2 I_m \quad (\text{III.5})$$

Avec :

β_1 et β_2 : les coefficients de régression associés à chaque variable indépendante

III.4.1-a Détermination des coefficients de corrélation

$$\beta = (X^T X)^{-1} X^T P_m = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} -5.9417 \\ 12.9102 \end{pmatrix} \quad (\text{III.6})$$

Avec :

X est la matrice des variables explicatives (avec éventuellement une colonne de 1 pour l'ordonnée à l'origine s'il y en a),

$$X = \begin{bmatrix} T_{a1} & I_{m1} \\ T_{a2} & I_{m2} \\ . & . \\ . & . \\ T_{21043} & I_{21043} \end{bmatrix}$$

P_m : Vecteur des observations de la puissance moyenne horaire,

β_1 et β_2 : Les coefficients de corrélation à estimer

L'analyse des coefficients de régression obtenus montre que la température ambiante T_a a un impact négatif sur la puissance prédite du parc photovoltaïque, avec un coefficient

de 5.9417. Cela signifie qu'une augmentation d'une unité de ces variables entraîne une diminution de la puissance produite.

à l'inverse, l'irradiation solaire I_m influence positivement la production d'énergie, avec un coefficient de 12.9102, indiquant qu'une augmentation d'une unité de l'irradiation se traduit par une augmentation significative de la puissance fournie.

Ces résultats sont cohérents avec le comportement physique attendu d'un système photovoltaïque, où l'irradiation est le principal moteur de la production d'énergie, tandis que des conditions de température ambiante élevée peut réduire le rendement des panneaux.

III.4.1-b Résultats et interprétation

Le calcul de R^2 (coefficient de détermination) est donné par :

$$R^2 = 1 - \frac{\sum_{i=1}^n \hat{E}_i^2}{\sum_{i=1}^n (P_{m_i} - \bar{P}_m)^2} = 1 - \frac{SCE_{reg}}{SCE_t} = 0.9557$$

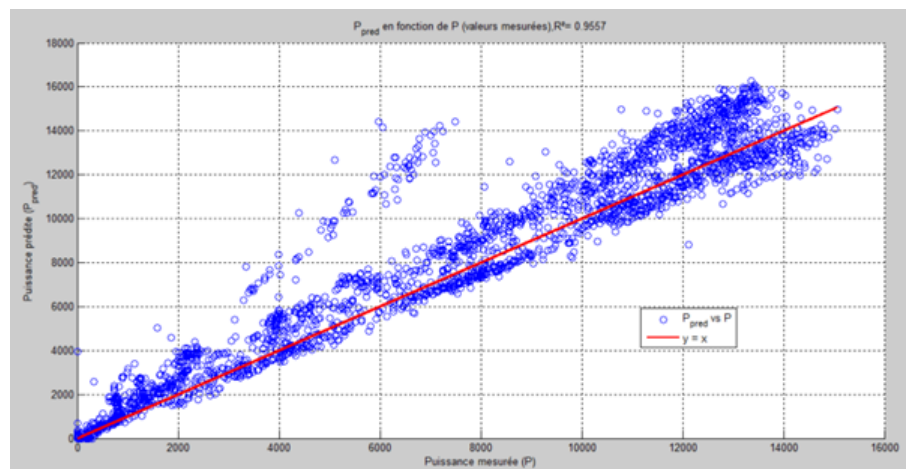


FIGURE III.6 – La corrélation entre puissance de sortie PV et la puissance prédit

La Figure III.6 montre la comparaison entre la puissance mesurée et la puissance prédite à l'aide de la régression linéaire multiple. Les points sont globalement bien alignés autour de la droite idéale $y=x$, indiquant une bonne performance du modèle. Le coefficient de détermination obtenu ($R^2 = 0.9557$) démontre que plus de 95 % de la variance de la puissance mesurée est expliquée par le modèle RLM. Toutefois, une légère dispersion est observée pour les valeurs élevées, traduisant certaines limites du modèle linéaire dans des conditions de forte irradiation.

1. Résultats de la Prédiction

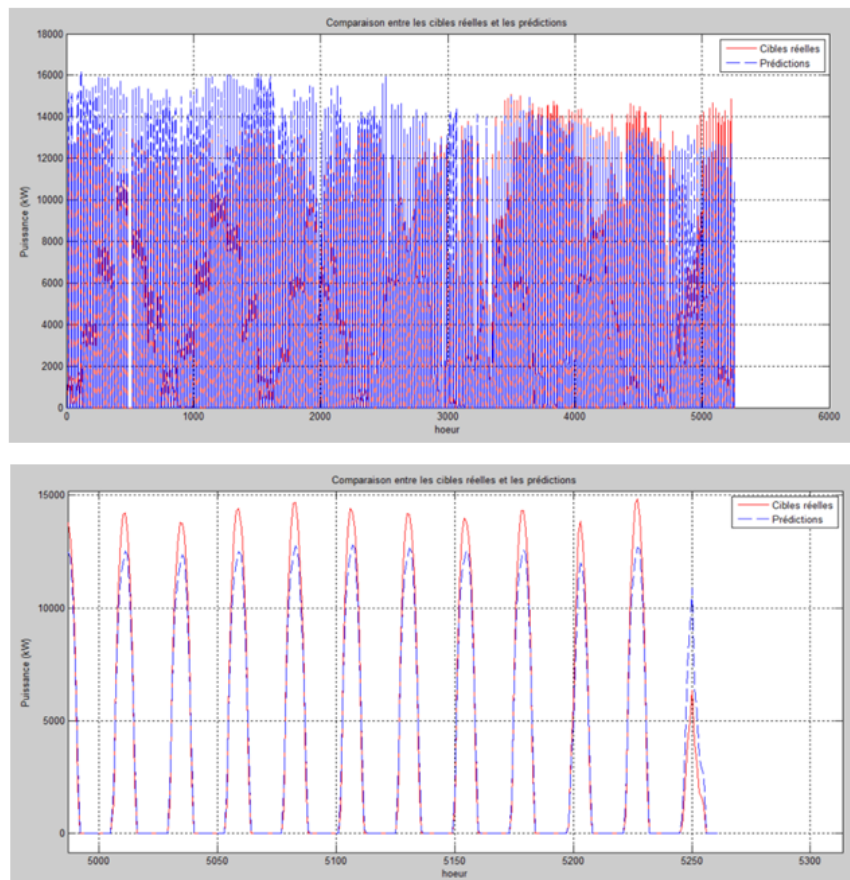


FIGURE III.7 – Représentation de la puissance de sortie PV et la puissance prédit (RLM)

La figure illustre la comparaison entre la puissance solaire mesurée et prédite en fonction du temps. Il est observé que la tendance générale de la production est bien suivie par la régression linéaire multiple, avec un alignement globalement satisfaisant des courbes. Néanmoins, quelques écarts sont visibles lors des périodes de haute production et de fortes variations, ce qui montre les limites du modèle linéaire pour capturer toute la dynamique non linéaire de la production solaire. Cette observation confirme la nécessité d'utiliser des approches plus avancées pour corriger ces écarts.

2. Analyse des Erreurs :

(a) Erreurs normalisées :

$$E = \frac{P_{m_i} - P_{prd_i}}{\text{Max}(P_{m_i})}$$

Erreur de biais moyenne normalisée (NMBE)

$$NMBE = \left(\frac{1}{M} \cdot \sum_{i=1}^M E_i \right) \times 100$$

Erreur absolue moyenne normalisée

$$NMAE = \left(\frac{1}{M} \cdot \sum_{i=1}^M |E_i| \right) \times 100$$

Erreur quadratique moyenne normalisée

$$NRMSE = \sqrt{\left(\frac{1}{M} \cdot \sum_{i=1}^M (E_i)^2 \right)} \times 100$$

écart type de l'erreur (SD)

$$SD = \sqrt{\frac{1}{M-1} \cdot \sum_{i=1}^M (E_i - \hat{E}_i)^2}$$

Où i est l'heure générique des données prédites est M est le nombre de données prédites, égal à 5261

(b) **Tableau des erreurs :**

TABLE III.5 – Tableau des erreurs de RLM

Les erreur	Les valeurs
NMBE	-2.2077 %
NMAE	3.3864 %
NRMSE	6.9308 %
SD	0.0657
R^2	0.9557

(c) **Interprétation :**

Le tableau présente les indicateurs d'erreur obtenus pour la régression linéaire multiple (RLM). Le biais moyen (NMBE) est faible et négatif (-2.21 %), traduisant une légère tendance à la sous-estimation des puissances solaires. L'erreur absolue moyenne (NMAE) est de 3.38 %, indiquant une bonne précision du modèle. Le NRMSE, égal à 6.93 %, montre que la dispersion des erreurs reste modérée par rapport à la valeur maximale observée. Enfin, l'écart-type des erreurs (SD = 0.0657) confirme une bonne stabilité du modèle. Ces résultats montrent que la RLM offre des performances satisfaisantes pour la prédiction globale de la puissance solaire, bien que certaines améliorations soient envisageables pour capter les dynamiques plus complexes.

III.4.2 Application de réseau de neurone

Dans cette section, nous proposons un modèle de prédiction basé sur un réseau de neurones, qui vise à estimer la puissance électrique générée à partir des données climatiques et environnementales mesurées. Le réseau sera entraîné sur un ensemble de données historiques, et sa performance sera évaluée en utilisant différents indicateurs d'erreur (NRMSE, MAE, R^2 , NMAE, NMBE, SD).

III.4.2-a Description de structure de réseau de neurone

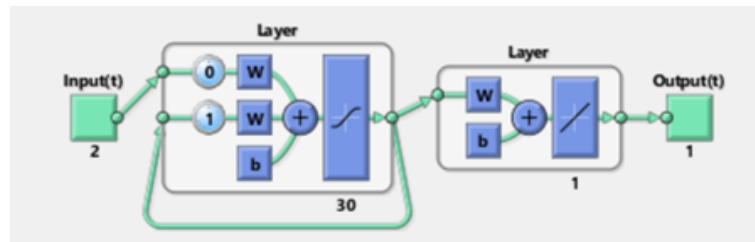


FIGURE III.8 – la structure de réseau de neurone

"Dans cette méthode, nous utilisons un réseau de neurones Elman (NAR) qui comprend deux entrées (température ambiante et irradiation) et une sortie (la puissance prédite)."

Première couche cachée (Layer) : Cette couche est composée de 30 neurones. Elle reçoit les entrées actuelles ainsi que la sortie de ses propres neurones à l'instant précédent (via la boucle de rétroaction), permettant ainsi de conserver une mémoire des séquences temporelles.

Boucle de rétroaction : Une connexion récurrente permet à la sortie de la première couche cachée d'être réinjectée comme entrée à cette même couche au pas de temps suivant. Ceci constitue la caractéristique principale d'un réseau d'Elman et lui confère sa capacité à traiter des données séquentielles.

Deuxième couche (Layer) : Le réseau comporte une deuxième couche, qui semble être la couche de sortie. Elle est constituée d'1 neurone.

Nombre de sorties (Output(t)) : Le réseau produit une sortie à chaque pas de temps 't', représentant la prédiction de la puissance fournie par la parcelle de panneau solaire."

III.4.2-b Paramètres du réseau Elman

TABLE III.6 – Paramètres réseau Elman

Paramètres	Entrer
Fonction d'entraînement	TRAINGDX
Fonction d'apprentissage adaptatif	LEARNGD
Fonction de performance	MSE
Nombre de couches :	4
Neurones (couche 1) – entrée :	3
Neurones (couche 2) : couche caché 1	30
Neurones (couche 3) : couche caché 2	1
Neurones (couche 3) – sortie :	1
Fonction d'activation : couche cachée	TANSIG
Fonction d'activation : couche de sortie	PURELIN
époques	340

III.4.2-c Résultats et interprétation

1. Résultats de la Prédiction

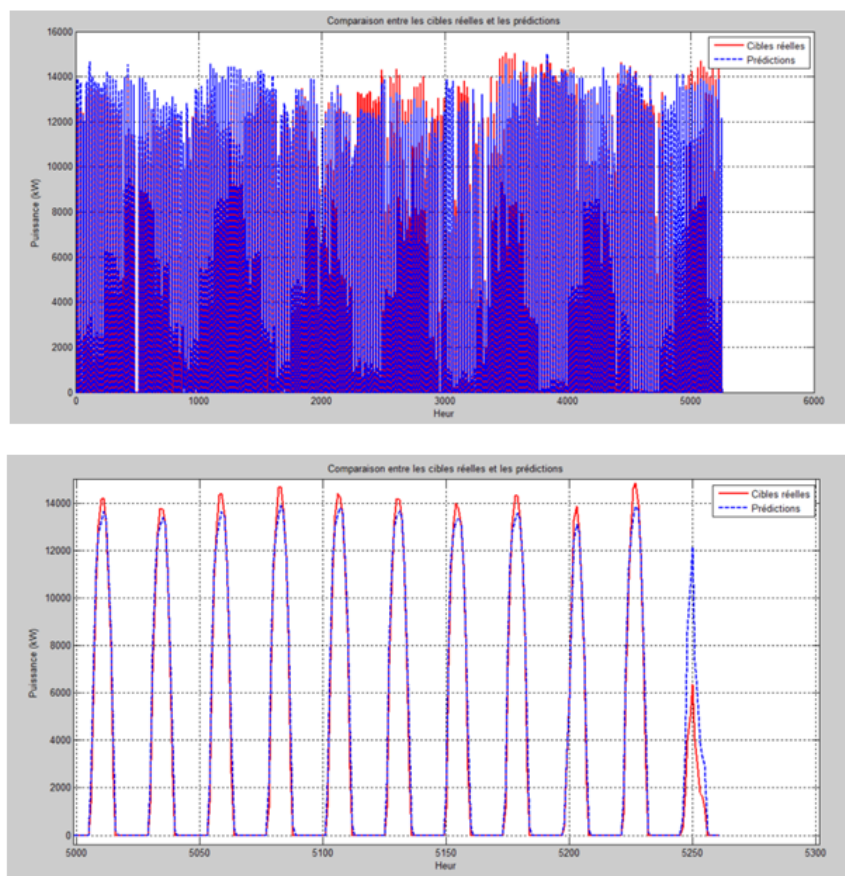


FIGURE III.9 – Représentation de la puissance de sortie PV et la puissance prédit (NAR)

(a) **Interprétation :**

"La représentation graphique de la comparaison entre la puissance réelle (courbe rouge) et les prédictions du réseau Elman (courbe bleue en pointillés) illustre visuellement la capacité du modèle à capturer la tendance générale et la cycli-
cité diurne de la production d'énergie solaire. On observe un suivi global des variations de puissance, avec les prédictions qui s'alignent sur les montées et les descentes de la courbe réelle.

Cependant, l'examen attentif révèle des écarts ponctuels entre les deux courbes, indiquant des erreurs de prédiction à des moments spécifiques. Ces écarts se manifestent par des sous-estimations (la courbe bleue est en dessous de la courbe rouge) et des sur-estimations (la courbe bleue est au-dessus de la courbe rouge).

(b) **Analyse des Erreurs**

TABLE III.7 – Tableau des erreurs de NAR

Les erreur	Les valeurs
NMBE	-1.2936 %
NMAE	2.3099 %
NRMSE	5.7635 %
SD	0.0562
R^2	0.9694

Le NMBE obtenu est de -1.2936 %, ce qui indique une légère sous-estimation systématique de la puissance prédite par rapport aux mesures réelles. Toutefois, cette valeur est faible, traduisant un modèle bien équilibré.

Le NMAE, égal à 2.3099 %, montre que l'erreur absolue moyenne est très faible, traduisant une bonne précision des prédictions.

Le NRMSE est de 5.7635 %, ce qui est inférieur à celui obtenu par la RLM seule, confirmant que le RNN améliore la qualité des prédictions en capturant mieux les dynamiques du système solaire.

L'écart-type ($SD = 0.0562$) montre que les erreurs de prédiction sont peu dispersées, assurant une bonne stabilité du modèle.

Enfin, le coefficient de détermination R^2 est de 0.9694, ce qui signifie que le modèle RNN explique près de 97 % de la variance observée dans les données réelles, montrant une excellente capacité prédictive.

Ces résultats démontrent que le modèle RNN est plus performant que la RLM seule pour la prédiction de la production photovoltaïque, notamment grâce à sa capacité à modéliser des relations non linéaires et temporelles complexes.

III.4.3 Régression Linéaire Multiple (RLM) combinée avec Réseau Elman (NAR)

Dans cette partie, nous proposons une approche hybride combinant la Régression Linéaire Multiple (RLM) et les Réseaux Elman (NAR). Le but est d'améliorer la précision de la prédiction de la puissance solaire en corrigeant les erreurs résiduelles de la RLM grâce aux capacités d'apprentissage séquentiel du NAR.

III.4.3-a Structuration de la méthode

La méthode hybride se déroule en deux grandes étapes successives :

1. Première étape : Régression Linéaire Multiple (RLM)

- **Entrées** : Température (T) et Irradiation (I).
- **Sortie** : Prédiction initiale de la puissance solaire

$$P_{prdit} = \beta_1 T_a + \beta_2 I_m$$

- **Principe** : La RLM cherche une fonction linéaire liant les entrées aux sorties :
- **Résultat** : La RLM fournit une prédiction rapide mais peut ne pas capturer toutes les dynamiques non linéaires du système.

2. Deuxième étape : Réseau Elman (NAR)

Les NAR sont capables de mémoriser des séquences temporelles et d'apprendre des comportements complexes à partir de données corrélées dans le temps. Ici, le NAR est utilisé pour prédire les résidus d'erreur de la RLM et ainsi affiner la prédiction initiale.

- **Entrées du NAR** : Les résidus d'erreur de la RLM, c'est-à-dire :

$$residus = P_{mesure} - P_{pred_{RLM}}$$

- **Sortie du NAR** : La correction à ajouter à la prédiction initiale.
- **Fonctionnement** : Le NAR apprend la structure temporelle cachée des erreurs. Il utilise sa mémoire interne pour modéliser les variations que la RLM n'a pas su apprendre.
- **Correction finale** :

$$P_{final} = P_{pred_{RLM}} + P_{corrig_{RNN}}$$

III.4.3-b Structuration globale du modèle

Le schéma général de l'approche hybride :

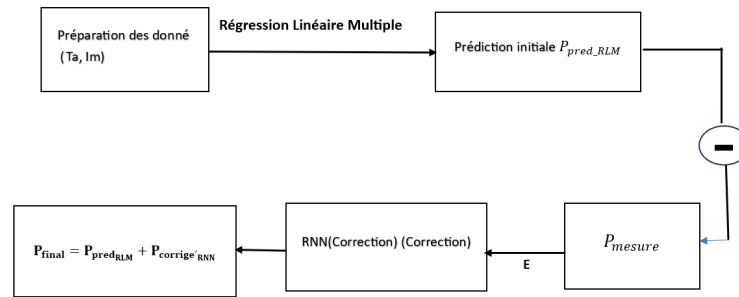


FIGURE III.10 – Structuration globale de RLM+RNN

E : Erreur résiduelle

Ainsi, RLM modélise les tendances linéaires et RNN modélise les erreurs complexes non linéaires.

III.4.3-c Résultats

1. Résultats de la Prédiction :

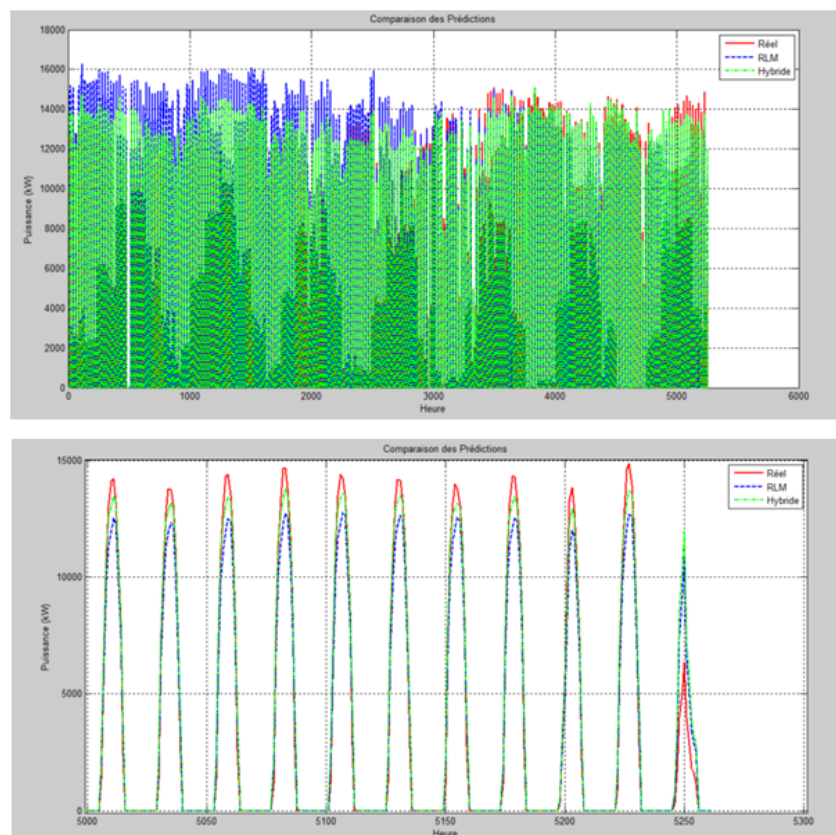


FIGURE III.11 – Représentation comparaison entre la puissance de sortie PV et la puissance prédit (methode de RLM+NAR)

Interprétation :

La figure montre clairement que la prédiction hybride (RLM+NAR) épouse beaucoup mieux la courbe réelle que la prédiction obtenue par RLM seul. Les écarts observés avec RLM, notamment dans les plages de haute et basse puissance, sont nettement réduits grâce à l'utilisation du NAR qui corrige les résidus du modèle linéaire. Ainsi, le modèle hybride améliore non seulement la précision mais aussi la stabilité des prédictions. Ces résultats confirment l'intérêt de la combinaison de modèles linéaires et neuronaux pour la prédiction de séries temporelles complexes.

2. Analyse des erreurs

TABLE III.8 – Tableau des erreurs de RLM+NAR

Les erreur	Les valeurs
NMBE	-1.14 %
NMAE	2.32 %
NRMSE	5.68 %
SD	0.0556
R^2	0.9703

3. Interprétation des erreurs :

NMBE = -1.14 % : Le Normalized Mean Bias Error est de -1.14 %. La valeur négative indique une légère tendance du modèle à sous-estimer la puissance fournie par le parc solaire. L'amplitude de -1.14 % suggère que cette sous-estimation est relativement faible par rapport à la production totale.

NMAE = 2.32 % : Le Normalized Mean Absolute Error est de 2.32 %. Cela signifie qu'en moyenne, les erreurs de prédiction de la puissance ont une magnitude de 2.32 % de la production totale. C'est une erreur moyenne relativement faible, indiquant une bonne précision globale en termes de taille des erreurs.

NRMSE = 5.68 % : Le Normalized Root Mean Square Error est de 5.68 %. Le NRMSE est toujours supérieur ou égal au NMAE. La différence entre les deux (5.68 % vs 2.32 %) suggère la présence de quelques erreurs de prédiction significativement plus importantes que la moyenne. La mise au carré des erreurs dans le calcul du RMSE donne plus de poids à ces erreurs plus importantes. Une valeur de 5.68 % reste dans une plage généralement considérée comme acceptable pour la prédiction de puissance solaire, mais indique une certaine dispersion des erreurs.

SD = 0.0556 : L'écart-type (Standard Deviation) des erreurs est de 0.0556. Cette valeur quantifie la dispersion des erreurs autour de leur moyenne (le biais, qui est reflété par le NMBE). Un écart-type de 0.0556 (ou 5.56 % si on le rapporte à une échelle normalisée similaire aux autres erreurs, bien que la normalisation ne soit pas explicitement indiquée ici) suggère une certaine variabilité dans la taille et la direction des erreurs. Combiné avec le faible NMBE, cela indique que si la tendance globale à la sous-estimation est faible, les erreurs individuelles fluctuent avec une dispersion

d'environ 5.56 %.

$R^2 = 0.9703$: Le coefficient de détermination (R^2) est de 0.9703. Cette valeur, très proche de 1, indique que le modèle RLM+RNN explique une très grande proportion (97.03 %) de la variance de la puissance fournie par le parc solaire. C'est un excellent R^2 , suggérant un très bon ajustement du modèle aux données et une forte capacité à prédire la puissance.

III.4.4 Régression Linéaire Multiple (RLM) combinée avec Filtre de Kalman

Pour surmonter les limitations de RLM, cette étude explore une approche hybride séquentielle combinant la RLM avec le filtre de Kalman, un estimateur récursif puissant pour le suivi des systèmes dynamiques en présence de bruit. Des approches similaires ont montré leur potentiel dans divers contextes de prédiction énergétique. Le filtre de Kalman corrige les erreurs de prédiction en temps réel. Les indicateurs de performance (NMBE, NMAE, NRMSE) montrent une nette amélioration par rapport à la RLM seule.

III.4.4-a Méthodologie

Notre méthodologie comprend deux étapes principales : l'application de la RLM sur les données d'entraînement pour établir un modèle initial, suivie de l'utilisation du filtre de Kalman sur les données de test pour améliorer les prédictions en temps réel.

1. Régression Linéaire Multiple (RLM) :

La RLM est une technique statistique utilisée pour modéliser la relation linéaire entre une variable dépendante et plusieurs variables indépendantes. Dans notre cas, la puissance solaire (P) est la variable dépendante, et l'irradiation solaire (I) et la température (T) sont les variables indépendantes. Le modèle est exprimé comme :

$$P_{prdit} = \beta_1 T_a + \beta_2 I_m$$

Les coefficients β_i sont estimés par la méthode des moindres carrés sur l'ensemble des données d'entraînement.

2. Filtre de Kalman pour l'Amélioration de la Prédiction :

Le filtre de Kalman est un algorithme d'estimation récursif qui fournit une estimation optimale de l'état d'un système dynamique en minimisant l'erreur quadratique moyenne. Il se compose de deux phases : la prédiction et la mise à jour. Dans notre application, nous adaptons le filtre pour affiner les prédictions initiales obtenues par la RLM. Un modèle d'état simple est considéré, où la puissance prédite par la RLM constitue l'état.

Le modèle d'état et le modèle de mesure sont définis comme suit :

$$P_{prdit_k} = AP_{prdit_{k-1}} + \omega_{k-1}$$

$$P_{mesure} = HP_{prdit_k} + \vartheta_k$$

Avec $A=1$ et $H=1$ dans notre implémentation, et $\omega_k \sim N(0, Q)$, $\vartheta_k \sim N(0, R)$ représentant le bruit de processus et de mesure, respectivement.

(a) **Implémentation Numérique :**

Le code MATLAB présenté précédemment est utilisé pour implémenter cette approche séquentielle. La RLM est d'abord appliquée pour obtenir les coefficients de régression (b), qui sont ensuite utilisés pour générer une prédiction initiale pour chaque point de données de test. Le filtre de Kalman utilise ces prédictions comme point de départ et les corrige en utilisant les mesures réelles de la puissance. (il nous reste un schéma explicative)

(b) **Choix des Paramètres du Filtre de Kalman**

Les paramètres Q (covariance du bruit de processus) et R (covariance du bruit de mesure) du filtre de Kalman jouent un rôle crucial dans ses performances. Q représente l'incertitude dans le modèle d'état, tandis que R représente l'incertitude dans les mesures. Dans cette implémentation, Q et R sont fixés à 5 et 1, respectivement. Ces valeurs ont été choisies empiriquement après plusieurs essais pour obtenir une bonne stabilité et une réponse adéquate du filtre. Il est important de noter que des méthodes plus avancées d'estimation des paramètres (comme celles discutées dans [64]) pourraient être utilisées pour optimiser ces valeurs.

(c) **Phase de Prédiction du Filtre de Kalman (écriture Théorique) :**

Dans le contexte de notre application, où nous utilisons un modèle d'état simplifié ($A=1$) et un modèle de mesure ($C=1$), l'étape de prédiction du filtre de Kalman peut être exprimée théoriquement comme suit :

- i. Prédiction de l'état (Puissance) : étant donné que notre modèle d'état suppose une évolution stationnaire ($A=1$), la prédiction de l'état à l'instant k basée sur l'estimation à l'instant $k-1$, notée $\hat{P}_{prdit_{kk-1}}$, est égale à la meilleure estimation de l'état à l'instant précédent, $\hat{P}_{prdit_{k-1k-1}}$, à laquelle on ajoute l'effet du bruit de processus. Cependant, dans notre implémentation spécifique, nous utilisons la prédiction issue du modèle de Régression Linéaire Multiple (RLM) à l'instant précédent comme point de départ pour l'estimation du filtre. Si nous devons formaliser cela dans le cadre du filtre de Kalman pur, et en considérant que la RLM nous fournit une estimation initiale de l'état, nous pourrions considérer que notre "état" est la puissance, et sa prédiction avant la mise à jour est influencée par la RLM.

Cependant, pour être précis avec les équations standard du filtre de Kalman appliquées à notre modèle simple, la prédiction de l'état serait :

$$\hat{P}_{prdit_{kk-1}} = A\hat{P}_{prdit_{k-1k-1}} = 1 \cdot \hat{P}_{prdit_{k-1k-1}}$$

Où $\hat{P}_{prdit_{k-1}}$ est l'estimation de la puissance à l'instant k1 après la mise à jour. Dans le code, $P_{est}(t1)$ joue ce rôle (pour t2). Pour la première étape du filtre (t=2), une "prédiction" initiale est en quelque sorte fournie par la RLM appliquée aux données de test à t1.

ii. Prédiction de la covariance de l'erreur de l'état :

La covariance de l'erreur de l'état prédite à l'instant k, notée P_{kk-1} est obtenue en propageant la covariance de l'erreur de l'état à l'instant k1, P_{k-1k-1} , à travers le modèle d'état et en ajoutant la covariance du bruit de processus Q_k (qui est constante et égale à Q dans notre cas) :

$$P_{k/k-1} = A_{k-1} \cdot \hat{P}_{k-1/k-1} \cdot A_{k-1}^T + Q_k = \hat{P}_{k-1/k-1} + Q_k$$

Dans notre code, $P_{cov_{pred}}$ à l'instant t correspond à $P_{t/t-1}$, et $P_{cov}(t1)$ correspond à $P_{t-1/t-1}$. Ainsi, la ligne de code.

III.4.4-b structurations de la méthode

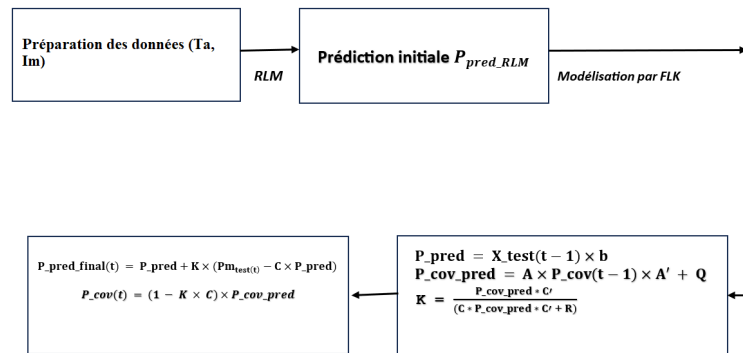


FIGURE III.12 – structuration de la méthode RLM + filtre de Kalman

FLK : Filtre de Kalman

X_{test} : Matrice des données de test

b : matrice de coefficient de corrélation

$P_{cov}(t)$: correspond à la covariance de l'erreur d'estimation à l'instant t.

Une faible valeur de $P_{cov}(t)$ indique une estimation plus fiable et vice versa

K(t) : le gain de Kalman

L'objectif du filtre de Kalman ici est de réduire les écarts résiduels entre la prédiction de la puissance P_{pred} et la puissance réellement mesurée $P_{m_{test}}$ en tenant compte dyna-

miquement de la confiance qu'on peut accorder au modèle par rapport aux données. Le résultat final est une estimation plus stable notée P_{est} , corrigée à chaque pas temporel.

III.4.4-c Résultats

1. Résultats de la Prédiction :

La figure ci-dessous illustre la comparaison entre la puissance réelle mesurée et la puissance estimée par le filtre de Kalman.

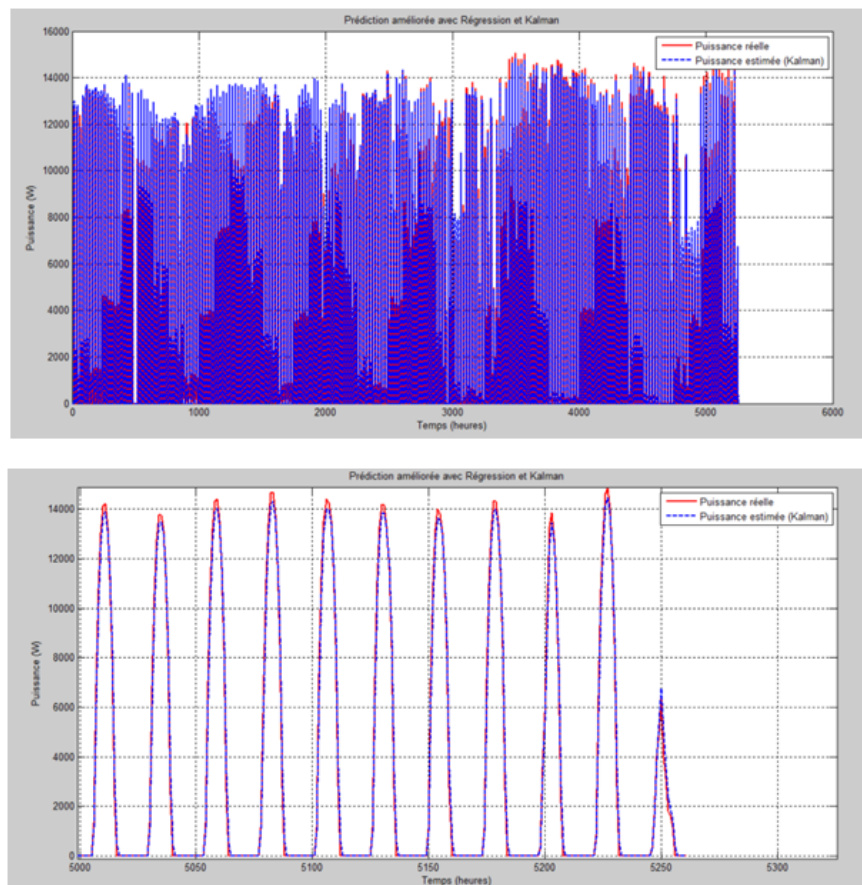


FIGURE III.13 – Représentation de la puissance de sortie PV et la puissance prédit (methode de RLM+filtre de Kalman)

Interprétation :

La figure montre une bonne concordance générale entre la puissance réelle et la puissance estimée. On observe que le filtre de Kalman parvient à suivre les variations de la puissance réelle, ce qui indique sa capacité à s'adapter aux changements dynamiques du système. Cependant, on note également quelques écarts, en particulier lors des pics et des creux, ce qui suggère que le modèle pourrait ne pas capturer parfaitement toutes les fluctuations rapides.

2. Analyse des Erreurs

Le tableau suivant présente les métriques d'erreur obtenues lors de l'évaluation du modèle combiné RLM-Filtre de Kalman :

TABLE III.9 – Tableau des erreurs de RLM+filtre de kalman

Les erreur	Les valeurs
NMBE	-0.3122 %
NMAE	1.2743 %
NRMSE	2.0881 %
SD	0.0206
R^2	0.9960

Interprétation des Erreurs

"L'évaluation de la performance du modèle RLM combiné avec le filtre de Kalman révèle les métriques d'erreur suivantes : NMBE (Normalized Mean Bias Error) = -0.1628 % : Cette valeur très faible et négative indique un biais global minime du modèle, avec une légère tendance à sous-estimer la variable prédite.

NMAE (Normalized Mean Absolute Error) = 1.4170 % : L'erreur absolue moyenne normalisée s'établit à 1.42 %, ce qui représente l'ampleur moyenne des erreurs de prédiction sans tenir compte de leur direction. Cela suggère une bonne précision globale du modèle.

NRMSE (Normalized Root Mean Squared Error) = 2.0931 % : L'erreur quadratique moyenne normalisée, qui est plus sensible aux erreurs importantes, est de 2.09 %. La légère différence entre le NRMSE et le NMAE suggère que les erreurs les plus importantes ne sont pas excessivement prononcées.

SD (Standard Deviation) = 0.0209 : L'écart-type des erreurs indique une faible dispersion des erreurs autour de leur moyenne. Cela suggère que les erreurs de prédiction sont relativement consistantes en magnitude.

En résumé, l'application du modèle RLM combiné avec le filtre de Kalman démontre une performance de prédiction très satisfaisante, caractérisée par un biais négligeable, une faible erreur moyenne et une dispersion des erreurs limitée. La sensibilité légèrement plus élevée du NRMSE par rapport au NMAE indique la présence de quelques erreurs plus importantes, mais dans l'ensemble, le modèle apparaît précis et fiable."

III.4.5 Comparaison entre les 4 méthodes

Le tableau ci-dessous montre les différents métriques d'erreur pour chaque que méthode.

Interprétation :

TABLE III.10 – Tableau des erreurs de différente méthode

Métrique d'erreur	Méthode RLM	Méthode NAR	Méthode RLM + Filtre de Kalman	Méthode RLM+NAR
NMBE	-1.1149 %	-1.3321 %	-0.1628 %	-1.11 %
NMAE	4.4127 %	2.4297 %	1.4170 %	2.30 %
NRMSE	6.9824 %	5.7192 %	2.0931 %	5.65 %
SD	0.0689	0.0556	0.0209	0.0554
R^2	0.9557	0.9699	0.9960	0.9703

Le tableau compare les performances de quatre méthodes de prédiction de puissance solaire : RLM, Réseau Elman, Filtre de Kalman, et RLM+RNN, en utilisant différentes métriques d'erreur. Analysons chaque métrique pour comparer les méthodes :

1. **NMBE (Normalized Mean Bias Error)** : Le Filtre de Kalman présente le biais le plus faible (-0.1628 %), indiquant la meilleure capacité à éviter une surestimation ou une sous-estimation systématique. RLM et Réseau Elman montrent des biais de sous-estimation plus importants (-1.1149 % et -1.3321 % respectivement). RLM+RNN se situe entre les deux, avec un biais de sous-estimation de -1.11 %. Conclusion (biais) : Le Filtre de Kalman est le plus performant en termes de réduction du biais.
2. **NMAE (Normalized Mean Absolute Error)** : Le Filtre de Kalman affiche l'erreur absolue moyenne la plus faible (1.4170 %), suggérant les prédictions les plus proches des valeurs réelles en moyenne. RLM est la méthode avec l'erreur absolue moyenne la plus élevée (4.4127 %). Le Réseau Elman (2.4297 %) et RLM+RNN (2.30 %) ont des erreurs moyennes similaires, légèrement supérieures à celle du Filtre de Kalman mais bien meilleures que RLM. Conclusion (magnitude de l'erreur) : Le Filtre de Kalman est le plus précis en termes de magnitude moyenne des erreurs, suivi de près par RLM+RNN et le Réseau Elman. RLM est la moins précise selon cette métrique.
3. **NRMSE (Normalized Root Mean Square Error)** : Le Filtre de Kalman présente l'erreur quadratique moyenne normalisée la plus faible (2.0931 %), indiquant une meilleure capacité à éviter les erreurs importantes. RLM a le NRMSE le plus élevé (6.9824 %), suggérant la présence d'erreurs significativement plus importantes que les autres méthodes. Le Réseau Elman (5.7192 %) et RLM+RNN (5.65 %) ont des NRMSE comparables, indiquant une présence similaire d'erreurs plus importantes, mais toujours mieux que RLM et moins bien que le Filtre de Kalman. Conclusion (sensibilité aux erreurs importantes) : Le Filtre de Kalman est le plus robuste face aux erreurs importantes, suivi par RLM+RNN et le Réseau Elman. RLM est le plus sensible aux erreurs importantes.
4. **SD (Standard Deviation)** : Le Filtre de Kalman a la plus faible dispersion des erreurs (0.0209), indiquant que les erreurs sont plus concentrées autour de la moyenne (le biais). RLM a la plus grande dispersion des erreurs (0.0689). Le Réseau Elman (0.0556) et RLM+RNN (0.0554) ont des dispersions d'erreurs similaires, légèrement inférieures à celle de RLM mais supérieures à celle du Filtre de Kalman. Conclusion (dispersion des erreurs) : Le Filtre de Kalman produit les erreurs les plus consistantes,

tandis que RLM a les erreurs les plus variables. Le Réseau Elman et RLM+RNN ont une variabilité d'erreur similaire.

5. R^2 (**Coefficient de Détermination**) : Le Filtre de Kalman affiche le R^2 le plus élevé (0.9960), indiquant la meilleure capacité à expliquer la variance des données de puissance solaire. Le Réseau Elman (0.9699) et RLM+RNN (0.9703) ont des R^2 élevés et similaires, suggérant une bonne capacité à capturer les tendances. RLM a le R^2 le plus faible (0.9557), bien que toujours élevé, indiquant une moins bonne explication de la variance par rapport aux autres méthodes. Conclusion (explication de la variance) : Le Filtre de Kalman est le modèle qui explique le mieux la variabilité de la puissance solaire, suivi de près par RLM+RNN et le Réseau Elman. RLM explique le moins bien la variance.

III.5 Conclusion

L'application de différentes méthodes statistiques pour la prédiction de la puissance solaire du parc étudié révèle des performances variables. Parmi les méthodes analysées, le (Filtre de Kalman+RLM) se distingue comme la méthode la plus performante sur plusieurs indicateurs clés : il présente le biais le plus faible, l'erreur absolue moyenne la plus basse, la plus faible sensibilité aux erreurs importantes et la plus faible dispersion des erreurs, tout en expliquant la plus grande partie de la variance des données.

Le RLM+RNN et le Réseau Elman montrent des performances globalement similaires, se situant entre le Filtre de Kalman et le RLM. Ils offrent une amélioration significative par rapport au RLM, qui apparaît comme la méthode la moins performante au regard des métriques d'erreur considérées.

Ces résultats soulignent l'importance du choix de la méthode de prédiction. Des techniques plus avancées, comme le (Filtre de Kalman+RLM) et potentiellement les approches hybrides (RLM+RNN), semblent mieux adaptées pour capturer la complexité et la variabilité de la puissance solaire. Bien que toutes les méthodes aient démontré une certaine capacité à prédire la puissance (avec des R^2 élevés), la précision et la fiabilité des prédictions varient considérablement.

Pour des applications où la précision est cruciale, notamment pour la gestion du réseau et l'optimisation de l'utilisation de l'énergie solaire, le (Filtre de Kalman+RLM) apparaît comme le choix le plus judicieux parmi les méthodes statiques étudiées. Des recherches futures pourraient explorer des hybridations plus poussées ou l'intégration de données météorologiques plus fines pour améliorer encore la précision des prédictions.

Conclusion générale

Dans un contexte mondial marqué par la transition énergétique et la lutte contre le changement climatique, le développement des énergies renouvelables représente un enjeu stratégique majeur. Parmi celles-ci, l'énergie solaire se distingue comme une source propre, durable et inépuisable à l'échelle humaine. Son exploitation constitue une réponse concrète aux défis énergétiques et environnementaux actuels.

Ce mémoire s'est intéressé à l'étude et à la prédiction de la production d'énergie photovoltaïque d'un parc solaire situé à Adrar, dans le sud-ouest de l'Algérie, une région caractérisée par un fort ensoleillement annuel. Ce travail a permis d'approfondir la compréhension du fonctionnement des systèmes photovoltaïques, des types de cellules solaires, ainsi que des principaux paramètres influent leur performance.

L'étude a reposé sur l'analyse de données environnementales et l'application de modèles de prédiction basés sur des approches statistiques et physiques. Ces méthodes permettent d'anticiper la production énergétique, d'optimiser l'exploitation des installations solaires, et de mieux intégrer cette énergie intermittente dans le réseau électrique. Une prédiction fiable contribue également à une meilleure planification des ressources et à une amélioration de la rentabilité des projets solaires.

Les résultats obtenus soulignent toutefois les limites de certains paramètres, comme l'humidité ou la température du module, qui présentent une faible corrélation avec la puissance produite. Cela met en évidence la nécessité d'enrichir les modèles en y intégrant des variables plus représentatives, telles que l'irradiation solaire directe, la nébulosité ou encore la poussière atmosphérique. L'utilisation de données météorologiques locales, précises et en temps réel, pourrait également améliorer significativement la qualité des prédictions.

En perspective, ce travail ouvre la voie à l'intégration de méthodes hybrides, alliant modèles physiques traditionnels et techniques d'intelligence artificielle avancées. En particulier, l'application d'approches telles que le Deep Learning, à travers des architectures comme les réseaux de neurones profonds (DNN) ou les modèles LSTM, permettrait de capturer des relations complexes et non linéaires entre les variables, et d'améliorer la robustesse et la précision des prévisions.

Ainsi, cette étude constitue une contribution à la valorisation du potentiel solaire algérien, notamment dans les zones sahariennes, et pose les bases pour le développement de solutions prédictives innovantes, mieux adaptées aux réalités climatiques locales et aux besoins du système énergétique de demain.

Bibliographie

Bibliographie

- [1] ADEME, Le photovoltaïque : principes et fonctionnement, Agence de la transition écologique, 2021.
- [2] ENEDIS, Raccordement d'une installation photovoltaïque au réseau public de distribution d'électricité, Guide technique, 2020.
- [3] ADEME, Les fondamentaux de l'énergie solaire photovoltaïque, Agence de la transition écologique, 2020.
- [4] HESPUL, Guide de l'énergie solaire photovoltaïque - Principe, installation et fonctionnement, Association Hespul, 2019.
- [5] <https://2u.pw/pDd9V>
- [6] <https://www.ecosources.org/types-de-cellules-photovoltaiquesutm-source=chatgpt.com>
- [7] <https://2u.pw/yQdIT>
- [8] <https://2u.pw/UH8ra>
- [9] <https://2u.pw/VlGhV>
- [10] <https://2u.pw/UzcXg>
- [11] BELAID LALOUNI Sofia « cours énergie solaire photovoltaïque », Université A.MIRA de BEJAIA. <https://biblio.univ-annaba.dz/ingeniorat/wp-content/uploads/2022/03/setar-memo.pdf>
- [12] M .BELHADJ ; Modélisation D'un Système De Captage Photovoltaïque Autonome Mémoire de Magister ; Centre Universitaire De Bechar ; 2008.
- [13] M. BENGHANEM. Etude de la performance des systèmes photovoltaïque autonomes Travaillant à poste fixe. Thèse de Doctorat, U.S.T.H.B, avril 1999.
- [14] B. EQUER. Energie solaire photovoltaïque. Volume I – Physique et technologie de la conversion photovoltaïque. CNRS. Edition ellipses, 1993.
- [15] M. BURESCH. Photovoltaic energy systems. McGraw-Hill U.S.A, 1983.
- [16] F. Lasnier, T. G. Ang, « Photovoltaic Engineering Handbook », IOP Publishing Ltd.1980.ISBN 0-85274-311-4.

- [17] E. BUCHET. Etude du dimensionnement et développement d'un logiciel d'aide à la conception de système de production d'énergie utilisant la conversion photovoltaïque de l'énergie solaire, Thèse de Doctorat, Faculté de Science et Technique de Saint-Jérôme, d'Aix Marseille, juillet 1988.
- [18] A. HADJ ARAB. Modélisation et simulation d'un système photovoltaïque de faible puissance, Thèse de magister, H.C.R, décembre 1989.
- [19] F. Lasnier, T. G. Ang, « Photovoltaic Engineering Handbook », IOP Publishing Ltd. 1980.ISBN 0-85274-311-4
- [20] <https://urls.fr/2tBnfR>
- [21] <https://2u.pw/sgDTV>
- [22] <https://2u.pw/mQpkQ>
- [23] <https://2u.pw/Lnl8V>
- [24] ADEME, Le photovoltaïque en France – Chiffres clés, Édition 2023.
- [25] Météo-France, Variabilité de l'ensoleillement et prévision solaire, 2021.
- [26] KETY, K., AMOU, K. A., SAGNA, K., et al. Modélisation et simulation d'un générateur photovoltaïque : Cas du module polycristallin Ecoline LX-260P installé au dispensaire de Sévagan (Togo). Journal of Renewable Energies, 2016, vol. 19, no 4, p. 633-645.
- [27] LA SALLE, Josselin Le Gal. Qualité et valeur des prévisions solaires probabilistes. 2021. Thèse de doctorat. Université de la Réunion.
- [28] <http://dspace.univ-tiaret.dz/bitstream/123456789/11834/5/CHAPITRE%20III.pdf>
- [29] <https://staff.univ-batna2.dz/sites/default/files/benderradji-ammar/files/cours-energies-renouvelables-suite.pdf>
- [30] <https://www.academia.edu/32811464/Principles-and-Performance-of-Solar-Energy-Thermal-Systems-A-Web-Course-by>
- [31] SMAIL, Mansouri, NOUREDINE, Pr Djarfour, MOHAMMED, Mr Boussaid, et al. Modélisation de gisement solaire par lois paramétrique dans la zone d'Adrar.
- [32] MEFTI, A. et BOUROUBI, M. Y. Estimation et cartographie de la composante globale du rayonnement solaire. Revue des Energies Renouvelables, 1999, p. 219-224.
- [33] OULIMAR, I., BOUCHOUICHA, K., et KHELIF, C. Modèle Radiométrique Adéquat pour Caractériser l'Apport Energétique Optimal sur Site d'Adrar. In : International Conference on Energy and Sustainable Development ICESD. 2013.
- [34] <https://www.researchgate.net/publication/258835812>
- [35] LIU, Benjamin YH. Daily insolation on surfaces tilted toward the equator. Journal of ASHRAE, 1961, p. 53-59.
- [36] HAY, John E. Calculation of monthly mean solar radiation for horizontal and inclined surfaces. Solar energy, 1979, vol. 23, no 4, p. 301-307.
- [37] Modeling of photovoltaic module performance based on datasheet parameters" par Skarlatos Dedoussis et al. (Energy Conversion and Management, Volume 103, 2015, Pages 810-820)

- [38] DIME, Joseph EVAGLE. Incidence du cout du risque de défaut sur les marges de taux des banques camerounaises.
- [39] BERTRAND, Frédéric. Compléments sur la régression linéaire simple Anova et inférence sur les paramètres.
- [40] LABROUSSE, Christian. Introduction à l'économétrie. (No Title), 1972.
- [41] RAKOTOMALALA, R. Économétrie-La régression linéaire simple et multiple. Université Lumière Lyon, 2011, vol. 2.
- [42] <https://2u.pw/78tZn>
- [43] <https://2u.pw/DdYp2>
- [44] Giraud, R., Chaix, N. (1989). Econométrie. Presses Universitaires de France (PUF)
- [45] WELCH, Greg, BISHOP, Gary, et al. An introduction to the Kalman filter. 1995.
- [46] Cours de réseaux de neurones artificiels/ / <https://bu.univhttp://www.univ-usto.dz/theses-en-ligne/doc-num.php?explnum-id=1702>
- [47] <https://2u.pw/pKI0e>
- [48] <https://2u.pw/DM7VS>
- [49] <https://abdelhamid-djeffal.net/web-documents/coursrna.pdf>
- [50] SID'ELVADEL, Mohamed Salem et ENCADRE PAR AOUCHE, Abdelaziz. MODELISATION MATHEMATIQUE DES SIGNAUX PHYSIOLOGIQUE : APPLICATION SUR L'ELECTROCARDIOGRAMME (ECG). 2021. Thèse de doctorat.
- [51] <https://2u.pw/Fr6Dv>
- [52] Rotella D, « Filtrage optimal », Ecole nationale d'ingénieurs de Tarbes,2004.
- [53] Naaman Bounar, « systèmes linéaires multivariables », Université Mohammed Seddiki ben Yahia- Jijel ,2001.
- [54] Sayah Abdelkader, « commande en réseau des systèmes retardes avec observateur linéaire et simulation sur un moteur électrique « DC » », Mémoire de magistère, Université d'Oran es senia,2012.
- [55] S. J. Julier and J. K. Uhlmann, « A New Extension of the Kalman Filter to Nonlinear Systems», In Proc. Of Aero Sense : The 11th Int. Symp, On Aerospace/Defence Sensing, Simulation and Controls, 1997.
- [56] JUTTEN, Christian. Filtrage linéaire optimal. Cours de cinquième année du département 3i Options Image et Signaux et Automatique. Université Joseph Fourier-Polytech'Grenoble. France, 2010.
- [57] A. Nabil « implantation d'une identification en temps réel de la machine asynchrone a cage sur le dsp tms320lf 2407a », Mémoire de magistère, Université de Mohamed Kheider biskra,2004.
- [58] Benzahia Y, « observateurs non linéaires : application sur un trms », Mémoire de master, université Mohamed Boudiaf - m'sila,2016.
- [59] Belmahdi. Fatiha « application du filtre de kalman pour le débruitage des signaux », Mémoire de magister, Université Mouloud Mammeri, Tizi Ouzou,2015.

- [60] K. Mohammed, « le filtre de kalman pour la prédiction multisite des apports liquides annuels du bassin versant Cheliff », Mémoire de master, Ministère de l'enseignement supérieur et de la recherche scientifique, 2017.
- [61] Z. Zheng, « commande à haute performance et sans capteur mécanique du moteur synchrone à aimants permanents », Thèse de doctorat, institut national polytechnique de Toulouse, 2008.
- [62] Wikipedia. Énergie en Algérie. <https://fr.wikipedia.org/wiki/Énergie-en-Algérie>
- [63] Global Energy Monitor. Adrar Solar Plant. <https://www.gem.wiki/Adrar-Solar-Plant>
- [64] Power Technology. Adrar Solar PV Park. <https://www.power-technology.com/data-insights/power-plant-profile-adrar-solar-pv-park-algeria>
- [65] Algeria360. Inauguration d'une centrale solaire à Adrar. <https://www.algerie360.com>
- [66] ScienceDirect. Performance assessment of PV power plants in Adrar. <https://www.sciencedirect.com/science/article/abs/pii/S0973082623001503>
- [67] <https://www.researchgate.net/publication/377666775>
- [68] HAKAM, Youness, AHESSAB, Hajar, GAGA, Ahmed, et al. Design and simulation of a 5 KW solar-powered hybrid electric vehicle charging station with a ANN-Kalman filter MPPT and MPC-based inverter control for reduced THD. *Scientific African*, 2025, p. e02563.
- [69] GABOITAOLELWE, Jwaone, ZUNGERU, Adamu Murtala, YAHYA, Abid, et al. Machine learning based solar photovoltaic power forecasting : A review and comparison. *IEEE Access*, 2023, vol. 11, p. 40820-40845.
- [70] MONTGOMERY, Douglas C., PECK, Elizabeth A., et VINING, G. Geoffrey. Introduction to linear regression analysis. John Wiley & Sons, 2021.
- [71] A. Mellit, "Artificial intelligence techniques for modeling and forecasting solar radiation data : A review," *Applied Energy*, 2009.
- [72] S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, 1997.
- [73] H. Abdelli, A. Mellit, "Predicting daily global solar radiation using hybrid Artificial Neural Network models," *Energy Conversion and Management*, 2016.
- [74] MARDIA, Kanti V., GOODALL, Colin, REDFERN, Edwin J., et al. The kriged Kalman filter. *Test*, 1998, vol. 7, p. 217-282.
- [75] WANG, Jianzhou et HU, Jianming. A robust combination approach for short-term wind speed forecasting and analysis—Combination of the ARIMA (Autoregressive Integrated Moving Average), ELM (Extreme Learning Machine), SVM (Support Vector Machine) and LSSVM (Least Square SVM) forecasts using a GPR (Gaussian Process Regression) model. *Energy*, 2015, vol. 93, p. 41-56.
- [76] XU, Longqin et LIU, Shuangyin. Study of short-term water quality prediction model based on wavelet neural network. *Mathematical and Computer Modelling*, 2013, vol. 58, no 3-4, p. 807-813.).

Résumer :

Dans ce travail, nous avons présenté les résultats d'une étude portant sur la prédiction de la production d'énergie photovoltaïque au niveau de la centrale solaire d'Adrar, située dans une région saharienne bénéficiant d'un fort ensoleillement. L'objectif principal a été de développer et d'évaluer différents modèles de prévision de la puissance produite, en tenant compte de l'impact des paramètres climatiques. Nous avons d'abord étudié le fonctionnement des systèmes solaires photovoltaïques, en détaillant les technologies des cellules solaires, leurs caractéristiques électriques, ainsi que les différents types d'installations. Ensuite, nous avons appliqué plusieurs approches de prédiction : des méthodes physiques basées sur l'irradiance solaire et des modèles statistiques incluant la régression linéaire, les réseaux de neurones et le filtre de Kalman. Les résultats ont montré que certains facteurs, comme l'humidité ou la température des modules, ont une influence limitée sur la production. En revanche, l'irradiation solaire directe s'est révélée être un paramètre déterminant. L'utilisation de méthodes hybrides, combinant modèles statistiques et intelligence artificielle, a permis d'améliorer la précision des prévisions. Ce travail ouvre ainsi des perspectives pour l'optimisation de la gestion des installations solaires dans les zones à fort potentiel, et souligne l'intérêt d'intégrer des techniques avancées de modélisation pour mieux anticiper la variabilité de la production énergétique.

Mots-clés (français) : Prédiction, production d'énergie photovoltaïque, centrale solaire d'Adrar, région saharienne, paramètres climatiques, irradiation solaire, systèmes photovoltaïques, technologies des cellules solaires, régression linéaire, réseaux de neurones artificiels, filtre de Kalman, méthodes hybrides, intelligence artificielle, modélisation avancée, optimisation de la production, variabilité énergétique, gestion des installations solaires.

ملخص

في هذا العمل، عرضنا نتائج دراسة حول التنبؤ بإنتاج الطاقة الكهروضوئية في محطة أدرار للطاقة الشمسية، الواقعة في منطقة صحراوية تتمتع بشمس ساطعة. كان الهدف الرئيسي هو تطوير وتقييم نماذج مختلفة للتنبؤ بالطاقة المنتجة، مع مراعاة تأثير العوامل المناخية.

لقد شرحنا أولاً كيفية عمل أنظمة الطاقة الشمسية الكهروضوئية، مع تفصيل تقنيات الخلايا الشمسية وخصائصها الكهربائية، بالإضافة إلى أنواع التركيبات المختلفة. ثم طبقنا عدة مناهج للتنبؤ: الأساليب الفيزيائية القائمة على الإشعاع الشمسي والنماذج الإحصائية بما في ذلك الانحدار الخطي والشبكات العصبية ومرشح كالمان.

أظهرت النتائج أن بعض العوامل، مثل الرطوبة أو درجة حرارة الوحدة، لها تأثير محدود على الإنتاج. من ناحية أخرى، ثبت أن الإشعاع الشمسي المباشر هو عامل حاسم. أتاح استخدام الأساليب الهجينة، التي تجمع بين النماذج الإحصائية والذكاء الاصطناعي، تحسين دقة التنبؤات. وبالتالي فإن هذا العمل يفتح آفاقاً لتحسين إدارة المنشآت الشمسية في المناطق ذات الإمكانيات العالية، ويسلط الضوء على الاهتمام بدمج تقنيات النمذجة المتقدمة للتنبؤ بشكل أفضل بتقلبات إنتاج الطاقة.

مفاتيح الكلمات (باللغة العربية)

التنبؤ، إنتاج الطاقة الكهروضوئية، المحطة الشمسية بأدرار، المنطقة الصحراوية، العوامل المناخية، الإشعاع الشمسي، الأنظمة الكهروضوئية، تقنيات الخلايا الشمسية، الانحدار الخطي،

الشبكات العصبية الاصطناعية، مرشح كالمان، الطرق الهجينة، الذكاء الاصطناعي، النمذجة المتقدمة، تحسين الإنتاج، تقلب الطاقة، إدارة المنشآت الشمسية.

Summarized :

In this work, we presented the results of a study on predicting photovoltaic energy production at the Adrar solar power plant, located in a Saharan region benefiting from strong sunlight. The main objective was to develop and evaluate different models for predicting the power output, taking into account the impact of climatic parameters. We first explained how photovoltaic solar systems work, detailing solar cell technologies, their electrical characteristics, and the different types of installations. We then applied several prediction approaches: physical methods based on solar irradiance and statistical models, including linear regression, neural networks, and the Kalman filter. The results showed that certain factors, such as humidity or module temperature, have a limited influence on production. However, direct solar irradiation proved to be a determining factor. The use of hybrid methods, combining statistical models and artificial intelligence, has improved forecast accuracy. This work thus opens up prospects for optimizing the management of solar installations in high-potential areas and highlights the value of integrating advanced modeling techniques to better anticipate the variability of energy production.

Keywords (English):

Prediction, photovoltaic energy production, Adrar solar power plant, Saharan region, climatic parameters, solar irradiation, photovoltaic systems, solar cell technologies, linear regression, artificial neural networks, Kalman filter, hybrid methods, artificial intelligence, advanced modeling, production optimization, energy variability, solar installation management.