Département
Génie Mécanique
Université Echahid Hamma Lakhdar - El-Oued

جامعة الشهيد حمّه لخضر - الوادي
UNIVERSITY OF ELOUED

Echahid Hamma Lakhdar University - El Oued
جامعة الشهيد حمة لخضر - الوادي
Faculty of Technology
Department of Mechanical Engineering
Speciality of Electromechanics

Master’s Thesis

Theme

Faults Diagnosis PV Panels Using Artificial Intelligence and Simulation

Prepared by: GUESSOUM NASRADDINE NEFZAOUA ELMAHDI

Board of Examiners:

Dr. REMHA SOUHAIB	Chairman
Dr. LAOUAMER MESBAH	Supervisor
Dr. BAGGAT FETEH	Examiner

Academic year 2025/2026

Dedication

I dedicate this work to my beloved parents, whose endless love, sacrifices, encouragement, and support have been the foundation of my success throughout my academic journey.

To my family, for their patience, understanding, and constant motivation during the completion of this work.

To my teachers and professors, whose knowledge and guidance have inspired me to pursue excellence and deepen my understanding of science and engineering.

To my friends and colleagues, who have shared this journey with me and provided support and encouragement along the way.

Finally, I dedicate this thesis to everyone who has contributed, directly or indirectly, to the accomplishment of this work.

Abstract:

Photovoltaic (PV) systems have become one of the most widely adopted renewable energy technologies due to their ability to generate clean and sustainable electrical energy. However, the performance and reliability of PV installations can be significantly affected by various faults and environmental conditions, leading to power losses, reduced efficiency, and increased maintenance expenses. Therefore, the development of effective fault diagnosis techniques is essential to ensure the reliable operation of photovoltaic systems.

This thesis presents a comprehensive framework for detecting and classifying photovoltaic faults through the integration of models, simulations, and AI techniques. First, the operating principles of photovoltaic systems and the most common fault types, including partial shading, increased series resistance (Rs), decreased shunt resistance (Rsh), and bypass diode failures, are investigated. Their impact on the electrical performance of PV modules is analysed using their current-voltage (I-V) and power-voltage (P-V) characteristics.

A detailed photovoltaic model is then developed in MATLAB/Simulink based on the single-diode equivalent circuit. Various fault conditions are introduced into the model to generate representative datasets under different operating scenarios. Several electrical and statistical features, including photovoltaic current, voltage, power, fill factor, irradiance, temperature, RMS value, mean value, energy, and I–V curve slope, are extracted and used for fault characterisation.

To automate the fault diagnosis process, artificial intelligence techniques such as artificial neural networks (ANN), convolutional neural networks (CNN), and support vector machines (SVM) are employed. The developed models are trained and tested to classify the operating condition of the PV system into five categories: normal operation, partial shading, increased series resistance, decreased shunt resistance, and bypass diode failure. The obtained results demonstrate the effectiveness of AI-based methods in achieving accurate and reliable fault detection and classification.

The proposed approach contributes to improving the monitoring, maintenance, reliability, and energy efficiency of photovoltaic systems and highlights the potential of artificial intelligence for advanced renewable energy applications.

Keywords: Photovoltaic Systems, Fault Diagnosis, MATLAB/Simulink, Artificial Intelligence, ANN, PCA, CNN, SVM, Partial Shading, Series Resistance, Shunt Resistance and Bypass Diode Failure.

Résumé :

Les systèmes photovoltaïques (PV) sont devenus l'une des technologies d'énergie renouvelable les plus répandues grâce à leur capacité à produire une énergie électrique propre et durable. Cependant, les performances et la fiabilité des installations PV peuvent être fortement affectées par divers défauts et conditions environnementales, entraînant des pertes de puissance, une baisse d'efficacité et une augmentation des coûts de maintenance. Par conséquent, le développement de techniques efficaces de diagnostic des défauts est essentiel pour garantir le fonctionnement fiable des systèmes photovoltaïques.

Cette thèse présente un cadre complet pour la détection et la classification des défauts photovoltaïques grâce à l'intégration de modèles, de simulations et de techniques d'intelligence artificielle. Dans un premier temps, les principes de fonctionnement des systèmes photovoltaïques et les types de défauts les plus courants, notamment l'ombrage partiel, l'augmentation de la résistance série (Rs), la diminution de la résistance parallèle (Rsh) et les défaillances des diodes de dérivation, sont étudiés. Leur impact sur les performances électriques des modules PV est analysé à l'aide de leurs caractéristiques courant-tension (I-V) et puissance-tension (P-V).

Un modèle photovoltaïque détaillé est ensuite développé sous MATLAB/Simulink, basé sur le circuit équivalent à une diode. Diverses conditions de défaut sont intégrées au modèle afin de générer des jeux de données représentatifs pour différents scénarios de fonctionnement. Plusieurs caractéristiques électriques et statistiques, telles que le courant photovoltaïque, la tension, la puissance, le facteur de remplissage, l'éclairement, la température, la valeur efficace (RMS), la valeur moyenne, l'énergie et la pente de la courbe I-V, sont extraites et utilisées pour la caractérisation des défauts.

Pour automatiser le processus de diagnostic des défauts, des techniques d'intelligence artificielle, notamment les réseaux de neurones artificiels (RNA), les réseaux de neurones convolutifs (RNC) et les machines à vecteurs de support (SVM), sont employées. Les modèles développés sont entraînés et testés afin de classer l'état de fonctionnement du système photovoltaïque en cinq catégories : fonctionnement normal, ombrage partiel, augmentation de la résistance série, diminution de la résistance shunt et défaillance de la diode de dérivation. Les résultats obtenus démontrent l'efficacité des méthodes basées sur l'IA pour une détection et une classification précise et fiable des défauts.

L'approche proposée contribue à améliorer la surveillance, la maintenance, la fiabilité et l'efficacité énergétique des systèmes photovoltaïques et souligne le potentiel de l'intelligence artificielle pour les applications avancées en matière d'énergies renouvelables.

Mots-clés : Systèmes photovoltaïques, diagnostic de pannes, MATLAB/Simulink, intelligence artificielle, ANN, PCA, CNN, SVM, ombrage partiel, résistance série, résistance shunt et défaillance de diode de dérivation.

الملخص:

تُعد الأنظمة الكهروضوئية (PVS) من أهم تقنيات الطاقة المتجددة المستخدمة في إنتاج الطاقة الكهربائية النظيفة والمستدامة. ومع ذلك، فإن أداء وموثوقية هذه الأنظمة قد يتأثران بالعديد من الأعطال والظروف التشغيلية والبيئية المختلفة، مما يؤدي إلى انخفاض القدرة المنتجة وتراجع الكفاءة وارتفاع تكاليف الصيانة. لذلك أصبح تطوير تقنيات فعالة للكشف عن الأعطال وتشخيصها ضرورة أساسية لضمان التشغيل الأمثل للأنظمة الكهروضوئية.

تهدف هذه الرسالة إلى تطوير إطار عمل متكامل للكشف عن أعطال الأنظمة الكهروضوئية وتصنيفها من خلال الجمع بين النمذجة والمحاكاة وتقنيات الذكاء الاصطناعي. في البداية، تم دراسة المبادئ الأساسية لعمل الأنظمة الكهروضوئية وتحليل أكثر الأعطال شيوعاً، والتي تشمل التظليل الجزئي، زيادة المقاومة التسلسلية(Rs) ، انخفاض المقاومة التفرع (Rsh)، وأعطال الصمام الثنائي التجاوزي .(Bypass Diode) كما تم تحليل تأثير هذه الأعطال على الخصائص الكهربائية للمنظومة من خلال منحنيات التيار-الجهد (I–V) والقدرة-الجهد.(P–V)

بعد ذلك، تم بناء نموذج تفصيلي للمنظومة الكهروضوئية باستخدام برنامج MATLAB/Simulink اعتماداً على نموذج الدايود الأحادي. وتم إدخال مجموعة من الأعطال المختلفة إلى النموذج بهدف توليد بيانات تمثل حالات تشغيل متعددة. كما تم استخراج مجموعة من الخصائص الكهربائية والإحصائية المهمة، مثل التيار والجهد والقدرة ومعامل الامتلاء (FF) والإشعاع الشمسي ودرجة الحرارة والقيمة الفعالة (RMS) والمتوسط الحسابي والطاقة وميل منحنى التيار-الجهد، وذلك لاستخدامها في عملية تشخيص الأعطال.

ولأتمتة عملية التشخيص، تم توظيف عدد من تقنيات الذكاء الاصطناعي، من بينها الشبكات العصبية الاصطناعية (ANN)، والشبكات العصبية الالتفافية(CNN) ، وآلة المتجهات الداعمة .(SVM) وقد تم تدريب هذه النماذج واختبارها من أجل تصنيف حالة النظام الكهروضوئي إلى خمس فئات مختلفة هي: التشغيل العادي، التظليل الجزئي، زيادة المقاومة التسلسلية، انخفاض المقاومة التفرع، وعطل الصمام الثنائي التجاوزي. وأظهرت النتائج المتحصل عليها قدرة تقنيات الذكاء الاصطناعي على تحقيق تشخيص دقيق وموثوق للأعطال.

تساهم المنهجية المقترحة في تحسين مراقبة وصيانة الأنظمة الكهروضوئية وزيادة موثوقيتها وكفاءتها في إنتاج الطاقة، كما تبرز الإمكانات الكبيرة للذكاء الاصطناعي في تطبيقات الطاقة المتجددة الحديثة.

الكلمات المفتاحية :الأنظمة الكهروضوئية، تشخيص الأعطال، MATLAB/Simulink، الذكاء الاصطناعي، الشبكات العصبية الاصطناعية، الشبكات العصبية الالتفافية، آلة المتجهات الداعمة، التظليل الجزئي، المقاومة التسلسلية، المقاومة التفرع، عطل الصمام الثنائي التجاوزي.

Contents

Chapter II: Modelling and Simulation of PV Panels Under Fault Conditions

Chapter III: Artificial Intelligence-Based Fault Diagnosis

Chapter IV: Results, Validation, and Discussion.

Figures list

Tables list

General Introduction:

Photovoltaic energy plays a pivotal role today in the production of electrical energy. But the environmental conditions and operational stresses strongly influence the performance and reliability of PV systems. During long-term operation, photovoltaic modules are exposed to factors such as temperature variations, dust accumulation, ageing, and partial shading. These conditions can lead to various faults, including bypass diode failures, increased series resistance, decreased shunt resistance, open-circuit faults, and short-circuit faults. Such faults reduce power generation and efficiency, accelerate system degradation, increase maintenance costs, and affect the economic viability of PV installations. To ensure reliable operation, effective monitoring and fault diagnosis methods. Instead of the old method of the diagnosis of the faults from the PV panels, the advances in artificial intelligence (AI) enable automated, fast, and accurate PV fault diagnosis, often outperforming traditional techniques such as visual inspection, electrical testing, and thermal imaging. Methods such as ANN, CNN, SVM, PCA, decision trees, random forests, and other machine/deep learning algorithms can classify PV faults using electrical, environmental, and statistical data.

The objective of this thesis is to investigate and develop an AI-based fault diagnosis framework for photovoltaic panels using simulation-generated data under various fault conditions. Different machine learning and deep learning techniques are evaluated and compared to determine their effectiveness in detecting and classifying PV faults.

This thesis is organized into four chapters. Chapter I presents the fundamentals of photovoltaic systems and the main fault mechanisms affecting PV panels. Chapter II focuses on the modelling and simulation of PV panels under normal and faulty operating conditions. Chapter III introduces the artificial intelligence techniques employed for fault diagnosis, including data preprocessing, feature extraction, and classification methods. Finally, Chapter IV presents the simulation results, performance validation, comparative analysis of the developed models, and a comprehensive discussion of the obtained findings.

Chapter I:

PV Fundamentals & Fault Mechanisms

I.1 Introduction:

Photovoltaics is the process of converting sunlight directly into electricity using solar cells, which are mainly based on semiconductors; they convert solar radiation into electrical energy. It is a type of renewable energy; the first practical photovoltaic devices were demonstrated in the 1950s. Solar cell research and development got its first big boost from the aerospace industry in the 1960s, which required a power supply separate from "grid" power for satellite applications, but when semiconductors were developed, they found large-scale applications in many fields such as agriculture, buildings, and industry. [1]

In the 1980s, research into silicon solar cells was successful, and solar cells began to increase in efficiency. In 1985, silicon solar cells reached the 20% efficiency milestone. Over the next decade, the solar cell industry experienced steady growth of between 15% and 20%, mainly boosted by the remote power supply market. The year 1997 had a growth rate of 38%. This growth has continued through the decades, and today solar cells are recognised not only as a means of providing electricity and a better quality of life to those who do not have access to the grid, but they are also a means of powering the grid, providing a significant fraction of grid power in larger locations today.

However, the performance and efficiency of PV systems are highly dependent on their reliability. During long-term operation, PV installations are exposed to various environmental and operational stresses, such as temperature fluctuations, partial shade, dust accumulation, ageing of components, and electrical disturbances. These conditions can cause a variety of faults, including short circuits and open circuits, which lead to the degradation of solar energy.

I.2 Fundamentals of Photovoltaic Energy Conversion:

Photovoltaic energy conversion is the process of transforming solar radiation directly into electrical energy through the photovoltaic effect. When sunlight strikes a photovoltaic (PV) cell, photons transfer their energy to semiconductor materials, generating electron–hole pairs. The internal electric field of the p–n junction separates these charges, producing a flow of electric current.

I.2.1 Solar Radiation and Spectral Characteristics:

Solar radiation is the electromagnetic energy emitted by the Sun and received by the Earth, serving as the primary energy source for photovoltaic systems. Its spectrum includes ultraviolet, visible, and infrared wavelengths, which directly influence the efficiency of solar energy conversion.

I.2.1.1 Solar irradiance components:

Solar radiation is the electromagnetic emissions from the sun that extend across the electromagnetic spectrum from the highly energetic x-ray region through the ultraviolet, visible, and IR portions of the spectrum to the far IR and radio regions. These emissions interact with Earth's own electromagnetic and atmospheric envelope, resulting in large variations in the magnitude of solar radiation available for conversion into other forms of useful energy. Most solar energy conversion systems utilise only part of the solar spectrum. Therefore, the radiation consists of different components. *Direct or beam radiation, diffuse radiation,* and *albedo radiation*.

- **Diffuse solar radiation**: It is the radiation that manages to reach the Earth's surface after going through numerous changes in its path, such as interactions with atmospheric gases.
- **Direct solar radiation**: This radiation goes through the atmosphere and reaches the Earth without spreading.
- **Reflected solar radiation**: This refers to the fraction of solar radiation that the Earth's surface reflects on its own. This takes place in a phenomenon known as the albedo effect.

***Solar Irradiance* (G):** Irradiated power density in (W/m^2) of (solar) radiation. An example is the solar constant, that is, the solar irradiance measured just outside the earth's atmosphere, which is roughly 1,361 W/m^2.

The spectral irradiance as a function of photon wavelength (or energy), denoted by *F*, is the most common way of characterising a light source. It gives the power density at a particular wavelength. The units of spectral irradiance are in $Wm^{-2}\mu m^{-1}$. The Wm^{-2} term is the power density at the wavelength $\lambda(\mu m)$. Therefore, the (m^{-2}) refers to the surface area of the light emitter, and the μm^{-1} refers to the wavelength of interest. In the analysis of solar cells, the photon flux is often needed as well as the spectral irradiance. The spectral irradiance can be determined from the photon flux by converting the photon flux at a given wavelength to Wm^{-2}.The result is then divided by the given wavelength, as shown in the equation below. [1]

$$F(\lambda) = \Phi E \frac{1}{\Delta\lambda} \text{ in SI units} \quad (I.1)$$

where in SI units:

F(λ) is the spectral irradiance in $Wm^{-2}\mu m^{-1}$;

Φ is the photon flux in # photons $m^{-2}sec^{-1}$;

E and *λ* are the energy and wavelength of the photon in joules and metres, respectively. The spectral irradiance is more commonly expressed in terms of wavelength so that the following is true:

$$F(\lambda) = \Phi\, q \frac{1.24}{\lambda(\mu m)} \frac{1}{\Delta\lambda(\mu m)} \tag{I.2}$$

Where:

F(λ) is the spectral irradiance in $Wm^{-2}\mu m^{-1}$.

Φ is the photon flux in photons $m^{-2}sec^{-1}$.

E and *λ* are the energy and wavelength of the photon in eV and μm, respectively.

q is a constant of 1.6×10^{-19}.

I.2.1.2 Air mass effect

The air mass (AM) is the path length that light takes through the atmosphere normalised to the shortest possible path length (that is, when the sun is directly overhead). The air mass quantifies the reduction in the power of light as it passes through the atmosphere and is absorbed by air and dust. The air mass is defined as

This depends primarily on the Sun's angular altitude h, using points O, A, and M, and this angle *h* is shown in Figure I.1. We can write the length of the Sun's path through the atmosphere:

$$OM = \frac{OA}{\sin(h)} \tag{I.3}$$

$$AM = \frac{OM}{\mathrm{OA}} = \frac{1}{\sin(h)} \tag{I.4}$$

In the expression AM, this OM/OA ratio is designated.

Examples

AM 0: By convention, this designates solar radiation outside the atmosphere.

AM 1: Sun at its zenith (at sea level)

AM 1.5: the sun at 41.8° is chosen as the reference for photovoltaics.

AM 2: Sun at 30°

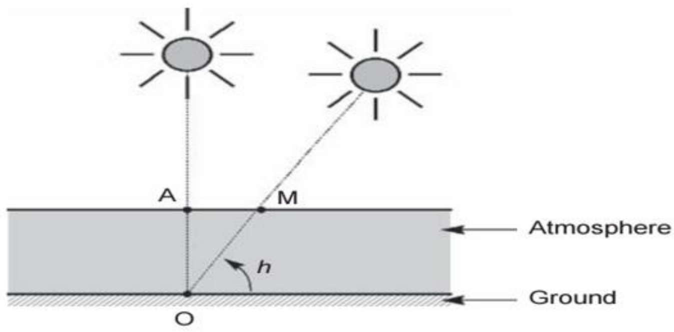


Figure I.1: Definition of Air Mass (AM).

I.2.1.3 Standard Test Conditions (STC):

The Standard Test Conditions (STC) are the backbone of solar panel performance ratings, or photovoltaic modules, that are of great importance in the development of solar energy technology. Due to the diverse environmental conditions to which solar panels are exposed, it becomes necessary to establish standardised performance indicator measurement standards, such as voltage, current, and power output; efficiency; and reporting. This highlights the importance of standard test conditions.

Standard test conditions refer to a set of controlled environmental parameters used during laboratory testing of solar panels. These conditions allow consistent performance comparisons across different PV technologies and manufacturers. According to the IEC 61215 and IEC 61730 standards, STC is defined by the following three parameters:

- Irradiance: 1000 W/m²
- Cell Temperature: 25°C
- Air Mass (AM) :1.5

In a nutshell, Standard Test Conditions are a fundamental pillar of the solar PV industry. They offer a unified and reliable basis for measuring, reporting, and comparing solar panel performance.

I.2.2 Photovoltaic Effect

I.2.2.1 Semiconductor physics:

The semiconductors are the basic component of solar cells, or photovoltaic (PV) panels. The semiconductors are able to absorb light (sunlight) and transfer some of the energy from the absorbed photons to electrons and holes, which are carriers of electrical current. Therefore, a solar cell is just a carefully engineered semiconductor, which can be fabricated from a number of semiconductor materials, for example, germanium (Ge), selenium (Se), cadmium telluride (CdTe), gallium arsenide (GaAs), gallium phosphide (GaP), sometimes with a small amount the copper indium Di selenide (CICS) or Cu(In, Ga)Se_2. The semiconductors are most commonly silicon (Si), crystalline, polycrystalline, and amorphous.

So, we must understand the basic semiconductors' physical principles underlying the operation of solar cells. A brief review of the fundamental properties of semiconductors is given that includes an overview of semiconductor carrier generation and band structure and the properties of the p-n junction generally. Next, carrier generation and carrier recombination, followed by a description of the basic operating characteristics of the solar cell.

Crystal structure:

By knowing the crystal structure of a material and its lattice dimensions, we can determine several characteristics of the crystal. For example, silicon is the semiconductor most widely used in the PV cell. In its eighth orbit from its atom, it has four electrons, which are called valence electrons. These electrons can form a chemical bond with a silicon atom, like it, in their crystal lattice. All of these bonds can be equal in length and angles; we can determine the volume density of atoms. The diamond lattice unit cell represents the real lattice structure of monocrystalline silicon. When in semiconductors, the concentration of impurity atoms is insignificant; we therefore call them 'intrinsic semiconductors.

Doping:

The doping is a method of manipulating the electrons and holes in crystalline silicon (C-Si) by replacing the silicon atoms in its crystal lattice with other atoms that contain five or three valence electrons, such as boron(B) and phosphorus(P), which are the most commonly used in crystalline silicon doping. So, in the process of doping a crystalline silicon lattice, four atoms are bonded to the five valence electrons of a phosphorus atom. When an electron absorbs a phosphorus atom at the thermal crystal lattice room temperature, the electron becomes free (moving); this process increases the concentration of electrons and is called 'impurity atoms', and donor atoms are symbolised (N_D).

But the process of doping of the boron (B), which contains three valence electrons, cannot bond with the four silicon atoms. However, it can easily receive an electron from the silicon bond. This is when thermal energy is formed in the silicon lattice chamber. Electron holes are formed that also receive impurity atoms, but the concentration of electron holes is called ‘acceptors’. We symbolise (N_A).

Carrier Transport in Semiconductors:

The properties of the transport in semiconductors, or the electrical conductivity of semiconductors, are moving charged particles with electric currents generated from the net flow of the holes and electrons. There are two basic mechanisms of transport in semiconductors: drift and diffusion.

- *Diffusion*: This is the natural tendency of particles to move from areas of high concentration to low concentration. Electrons diffuse from the n-side to the p-side.
- *Drift:* This is the movement of charge caused by an electric field. In a solar cell, the internal electric field at the junction "sweeps" the light-generated electrons back to the n-side.

In the dark, these two forces balance out (net current = 0). When sunlight hits, the balance is broken, and *Drift* takes over to push current through your device.

I.2.2.2 P–N junction behaviour:

The p-n junction behaviour brings us closer to understanding the principle of operation of solar cells, the n-p junction – a contact between a layer of p-type and a layer of n-type semiconductors, if they are the same base material, Si crystal structure, from the n-p junction. It’s called a “homogeneous junction”, but if they are different base materials or doped between the silicon and phosphorus or boron, it is known as a “heterogeneous junction”. On the other hand, an n-layer has an abundance of free electrons, and the p-layer has an abundance of free holes. Under thermal equilibrium conditions, meaning that temperature is the only external variable influencing the populations of free holes and electrons, the relationship between hole density, p, and electron density, n, as shown in the following figure (I.2) at any given point in the material, is given by the following:

$$np=n_i^2 \tag{5.I}$$

Where: n_i: the density of holes or electrons in intrinsic (impurity-free) material, approximately, $n \cong N_D$ and $p \cong N_A$ in impurities material; N_D: the density of donor impurity. N_A: the density acceptor impurity.

Again, we have summarised the different types of n-p junctions, as they are crucial to the semiconductor's function physically from the solar cell's operation.

- ✓ p-n homojunction that is in the same material as the semiconductor.
- ✓ p-n heterojunction is formed by two chemically different semiconductors.
- ✓ p-i-n junction is an addition of an essential layer (i) between the n-layer and p-layer, which in turn extends the electric field formed by the p-n junction.
- ✓ MS junction is the junction between a metal-type layer and semiconductor layers; an example is the Schottky barrier formed.

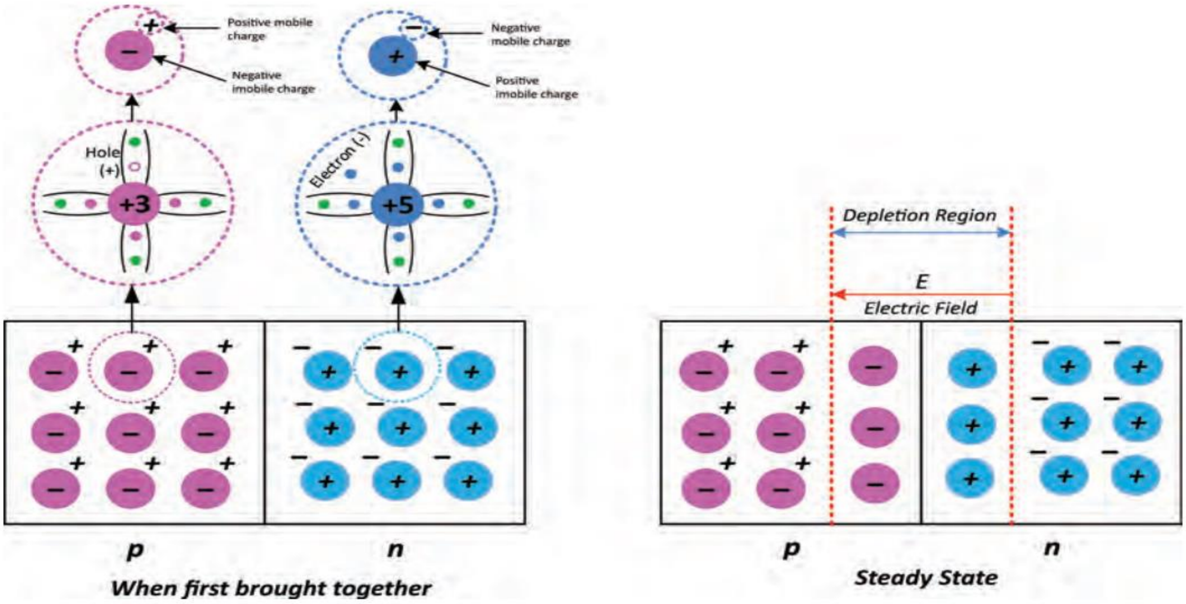


Figure I.2: This diagram illustrated the behaviour of the n-p junction and the formation of the electric field and depletion region from the semiconductors. [8]

1.2.2.3 Carrier generation and recombination:

When the semiconductors are exposed to the external phonon field, an electric field, or an electromagnetic field, the electron is excited from the valence band to the conduction band of the semiconductor, thereby creating a hole in the valence band. This phenomenon is called 'carrier generation' and is represented by the generation rate (G_{th}). But the opposite of this process is called recombination of carriers'; it is represented by the recombination rate (R_{th}). The recombination is the return of concentrations of electrons and holes toward the thermal equilibrium state, and any transfer of an electron from a conduction band to a valence band and the valence-band hole it's removed from is illustrated in Figure I. 3.

The mechanism of generation and recombination is important to the operation of solar cells, and among the multiple processes from the generation and recombination mechanism, the most significant are direct and indirect generation and recombination (band-to-band), and everyone has their own process. The generation process is signified by the photogeneration, but the recombination process is signified by recombination through traps (defects) in the forbidden gap, radiative recombination, and Auger recombination, as illustrated in Figure I. 3.

Figure I.3: the generation and recombination of electron-hole pairs.

The direct generation and recombination (band-to-band): theoretically, this process is one of the simplest of generation and recombination processes. In the generation process, a pair is free from electrons and holes, i.e., the conduction band has accepted an electron from the valence band. So again, it is with the recombination process that it creates a pair of a free electron and hole, i.e., the electron moves from the conduction band to the valence band. In that case, the semiconductors in thermal equilibrium and the carrier concentration remain constant with the condition maintained ($pn=n^2_i$). The generation and recombination rates are equal, as shown in figure I.4(a), so we have the following:

$$hv = E_2 - E_1 \tag{I.6}$$

$$G_{n0}=G_{p0}=R_{n0}=R_{p0} \tag{I.7}$$

$$G_{th} = R_{th} \tag{I.8}$$

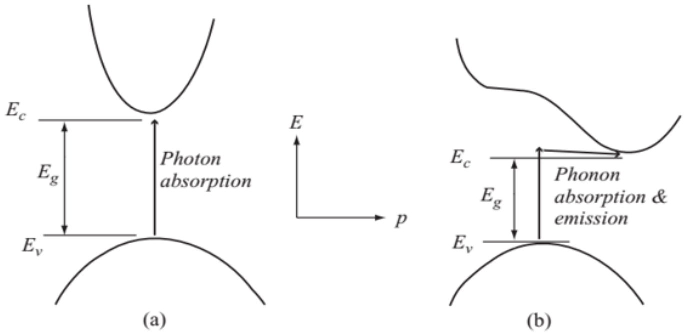


Figure I.4: diagram illustrating direct and indirect G-R

In indirect generation and recombination (indirect band gap) in semiconductors from two different atoms like Si and Ge, the maximum and minimum of the energy versus momentum relationship occur at a different crystal momentum. An electron transition between the maximum of the valence band and the minimum of the conduction band is not possible with only the absorption of a photon because the momentum of the photon is too small, as shown in figure I.4(b).

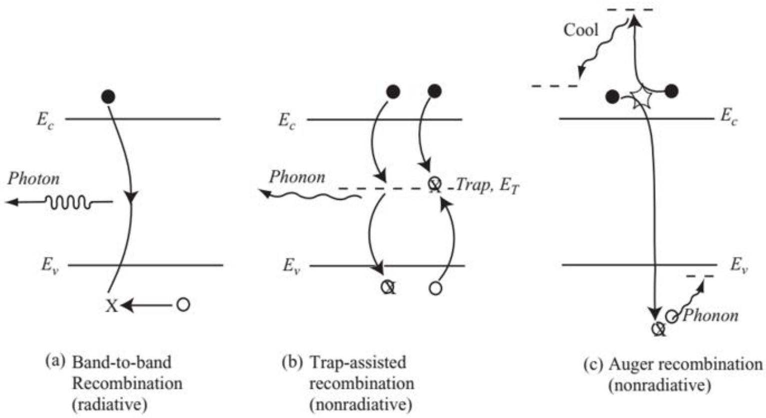


Figure I.5: The diagram illustrated the recombination process.

I.2.3 Structure of a Photovoltaic Cell:

I.2.3.1 The PV cell:

The photovoltaic (PV) cells convert solar radiation that is falling on the PV cell directly into electricity in a clean and environmentally friendly way. The PV cells are made of the semiconductor from the layer's material, one of which is negatively charged and the other is positively charged, which are called 'p-n junctions'. Once the sunlight reaches the PV cell, the semiconductor materials absorb some of the photons; they are generating the electron-hole pairs. This flow produces an electric current. The photovoltaic cells are laminated with various layers to protect them from environmental influences. There is an anti-reflecting coating layer is applied to the surface to increase photon absorption and reduce optical losses. The front contact and emitter, base and rear contact show to figures I.6.[14]

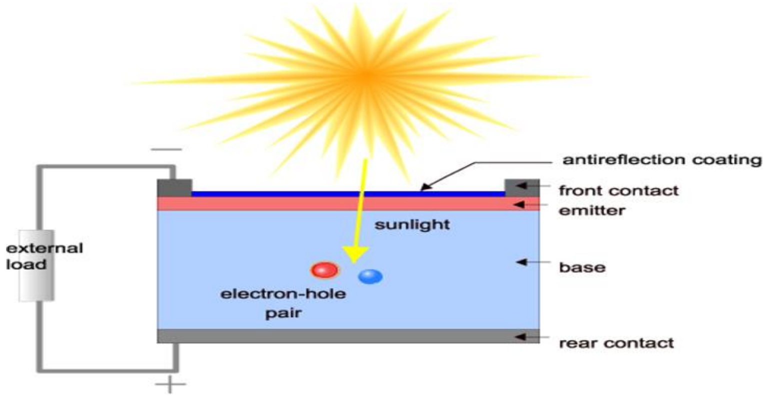


Figure I.6: The diagram illustrated the solar cell.

I.2.3.2 Anti-reflective coating:

The anti-reflective coating or (ARC) of the solar cells is a thin layer of dielectric material is applied on the surface of PV solar cells to minimise the photon absorption and light reflection into the semiconductors' solar cells. The (ARC) reduces the reflects silicon of incident sunlight to <2-5% via the cancellation of light waves reflected off the bottom and top surfaces of coating. It works to increase overall efficiency and boosts short-circuit current density (J_{SC}) by (20-30%) of solar cell. The thickness of AR coating is defined by the one-fourth of the beam wavelength incidental(λ/4),[13], and the materials used in the ARC like TiO_2, Si_3N_4 or SiO_2. The anti-reflective coating has different types. A single-layer anti-reflective coating (SLARC), double-layer anti-reflective coating (DLARC) and multi-layer anti-reflective coating. The thickness $d_1$of anti-reflective coating and refractive index n_1 represented by the equation [15]:

$$d_1 = \frac{\lambda_0}{4_{n_1}} \tag{I.9}$$

I.2.3.3 Emitter and base layers:

The emitter and base layers are the functions of each of the p-n-type layers doped in the solar cells.

The emitter layer in a crystalline silicon solar cell is an n-type doped layer; the density of donors of the emitter is high, and the diffusion potential from the p-n junction has to be a maximum. The thickness of the emitter (d_{em}) in the maximum density donor should be of the order of 100 nm and can be obtained.

$$d_{em} \leq \sqrt{\frac{D_p}{100.C_A.n_0^2}} \tag{I.10}$$

The base layer is the thick light-absorbing layer of the solar cell; this layer helps in creating an electric field. The base is wafering the doped definition, and the lifetime of the base layer in minority charge carriers is the minimum lifetime, about 100 µs, for example. Therefore, the base of a crystalline silicon solar cell is p-type doped with a density acceptor. The thickness of the base layer has to be of the order of 200-400 µm. [15]

I.2.3.4 Front and back contacts:

Contacts are the metallic electrodes that collect the separated charges and deliver them to an external circuit.

A. ***Front contacts****:* The front contact of the solar cell is narrow conductors that are typically connected by bus bars. Its width is a maximum of 200 µm to avoid light loss, and the distance between the front contacts determines the sheet resistance of the n^+ layer. The sheet resistance depends upon the phosphorus doping level. Which in turn is involved in the process of light absorption and recombination. This is defined by the sheet resistance Rs of the layer, expressed in units of ohms, where: Rs=$\frac{\rho}{T}$, (ρ) is the layer resistivity, and (T) is layer thickness. In order to ensure good ohmic contact, the silver-based paste is n-type doped with phosphorus [18, 19].

B. ***Back contacts****:* The back contact is usually composed of highly conductive metals, like an aluminium metal of p^+ doping that forms a layer. serves two functions. It creates the back surface field that has a low velocity in the recombination process, and it also allows for a tunnelling-type ohmic contact to the back aluminium metal contact. Commercially available aluminium pastes are available that contain the aluminium in a paste suspension. A low-temperature baking vaporises unwanted components and then the aluminium into the silicon through a high-temperature annealing process. A silver paste containing a small amount of aluminium is also available, which forms a silver metallic backing layer, allowing the back of the soldered module during manufacturing. The screen-printing method allows the silver paste to be deposited selectively onto small areas of the back of the wafer, which is done to enable soldering while minimising the amount of silver required. If the screen is only porous in the areas in which a deposit is desired, then when the rubber blade is passed over the screen during printing, paste is only applied to the silicon in these areas. Screens may be prepared with a spatially patterned polymer masking layer that blocks the pores of the screen in selected areas. A suitably masked screen may be used to print a given pattern onto thousands of wafers. [16]

I.3 Electrical Modelling of PV Cells and Modules

I.3.1 Ideal PV Cell Model:

The ideal circuit model for a PV cell is illustrated in Figure I.7. It is the simplest circuit, consisting of an ideal current source as the photocurrent (I_{ph}) in parallel with a real diode as the current source in the model, providing current in proportion to incident solar radiation on the cell surface.

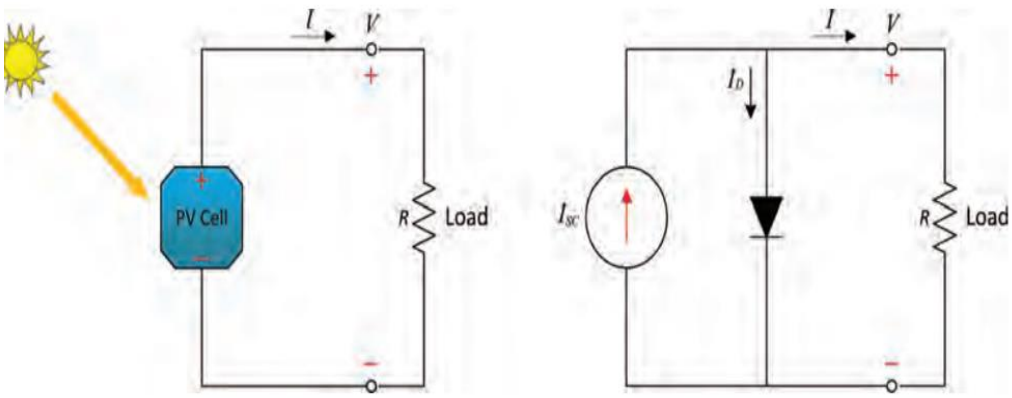


Figure I.7: ideal circuit model for a PV cell.

I.3.2 Single-Diode Model (SDM):

A model to represent crystalline-based PV cells is usually formed from the equivalent circuits. These models are usually categorised into two main types: single-diode models (SDMs) and double-diode models (DDMs).

The single-diode model is an ideal model for PV cells, which should have a current source in parallel with a diode. However, due to various non-ideal factors, the equivalent circuit of the standard SDM also includes one shunt resistor and one series resistor, as illustrated in Figure I.8. The current-voltage characteristics according to the equivalent circuit are expressed as:

$$I_{pv} = I_{ph} - I_d - \frac{V_d}{R_{sh}} \quad (I.11)$$

$$I_d = I_s\left[e^{\left(\frac{qv_d}{kT_CA_n}\right)} - 1\right] \tag{I.12}$$

where V_d is equal to

$$V_d = V_{pv} + I_{pv}\mathrm{R_s} \tag{I.13}$$

$$I_{pv} = I_{ph} - I_s\left[exp\left(\frac{V_{pv}+IR_S}{V_d}\right) - 1\right] - \frac{V_{pv+I.R_S}}{R_{sh}} \tag{I.14}$$

I_{pv}: Current supplied by the cell [A].

V_{pv} : Voltage at the cell terminal [V].

$I_{\ ph}$: The photocurrent, proportional to the irradiance.

I_s: The current saturation of the diode.

R_s : Series resistance [ohm].

R_{sh}: Shunt resistance (or parallel) resistor [ohm]

E_{STC}: Irradiance at STC 1000 W/m²

k: Boltzmann constant1.38 $\times 10^{-23}$ J/K

q: Charge 1.6 × 10^{-19} C

T_{CS}: PV cell temperature at STC 298 K or 25 C.

V_{TCS}: Thermal voltage of p-n junction at STC* 25.7 MV.

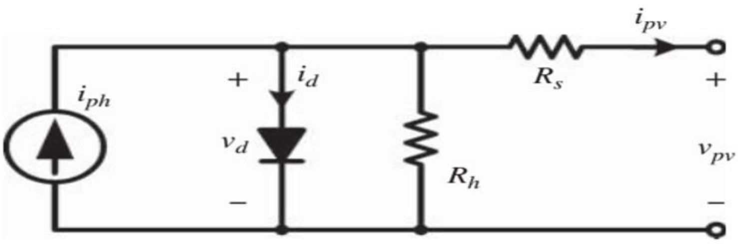


Figure I.8: Equivalent circuit single-diode model.

I.3.3 Double-Diode Model (DDM):

In figure 1.9 illustrated is the equivalent circuit, the dual-diode model. The DDM is an ideal model for PV cells; it should have a current source in parallel with a diode. Includes another diode in parallel and one shunt resistor (Rsh) and one series resistor (Rs). the equivalent circuit of the standard (SDM + diode) The current-voltage characteristics according to the equivalent circuit are expressed as:

$$I_{pv} = I_{ph} - I_{d1} - I_{d2} - \mathrm{I_{sh}} \tag{I.15}$$

$$I_d = I_s\left[e^{\left(\frac{qv_d}{kT_CA_n}\right)} - 1\right] \quad (I.16)$$

where V_d is equal

$$V_d = V_{pv} + I_{pv}R_s \quad (I.17)$$

$$I_{pv} = I_{ph} - I_{s1}\left[exp\left(\frac{q(V_{pv}+IR_s)}{n_1K_bT}\right) - 1\right] - I_{s2}\left[exp\left(\frac{q(V_{pv}+IR_s)}{n_2K_bT}\right) - 1\right] - \frac{V_{pv+I.R_S}}{R_{sh}}$$

(I.18)

I_{pv}: current provided by the cell [A].

V_{pv}: the voltage of the terminal of the cell [V].

I_{ph}: The photocurrent, proportional to the irradiance.

I_{s1} : corresponding to the current of saturation de la diode1.

I_{s2}: corresponding to the current of saturation current of diode 2,

R_s : Resistance series [ohm].

R_{sh}: Resistance shunt (or parallel) resistor [ohm]

$n1$, the non-ideality factor of the diode junction. [1].

$n2$, the non-ideality factor of the junction of diode 2,

E_{STC}: Irradiance at STC 1000 W/m²

k: Boltzmann constant1.38 $\times 10^{-23}$ J/K

q: Charge 1.6 × 10^{-19} C

T_{CS}: PV cell temperature at STC 298 K or 25 C

V_{TCS}: Thermal voltage of p-n junction at STC* 25.7 Mv.

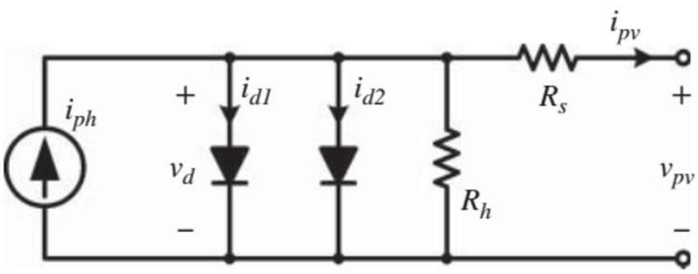


Figure I.9: Equivalent circuit of double-diode model.

I.3.3.1 Improved Recombination Modelling: The primary reason for using the double-diode model is to better represent recombination.

- In SDM: Recombination is simplified into a single diode, which is often inaccurate when the sunlight is weak.
- In DDM: A second diode is added specifically to model recombination losses in the depletion region of the solar cell. This makes the model much more accurate, especially at low solar radiation levels.

I.3.3.2 The Comparison: SDM vs. DDM: The comparison lies between the single diode model to the double diode model illustrated in table I.1.

Table I.1: SDM vs DDM.

Feature	**Single-Diode Model (SDM)**	**Double-Diode Model (DDM)**
Circuit Equivalent	A current source, **one** diode, and two resistors (Rs and Rsh).	A current source, **two** diodes in parallel, and two resistors (Rs and Rsh).
Equation	$I_{pv} = I_{ph} - I_s\left[exp\left(\frac{V_{pv}+IR_s}{V_d}\right) - 1\right]$- $\frac{V_{pv+I.R_S}}{R_{sh}}$	$I_{pv} = I_{ph} - I_{s1}\left[exp\left(\frac{q(V_{pv}+IR_s)}{n_1 K_b T}\right) - 1\right] I_{s2}\left[exp\left(\frac{q(V_{pv}+IR_s)}{n_2 K_b T}\right) - 1\right]$- $\frac{V_{pv+I.R_S}}{R_{sh}}$
Number of Parameters	**5 Parameters**: I_{pv}, I_s, a, R_s, R_{sh}	**7 Parameters**: I_{pv}, I_{S1}, I_{S2}, a_1, a_2, R_s, R_{sh}
Accuracy	Good for general use and fast calculations.	**Superior accuracy**, particularly for modeling temperature and low-light effects.
Complexity	Simple and easy to solve mathematically.	Complex; requires advanced algorithms to find the correct parameter values.

I.3.4 PV Module and Array Configuration:

I.3.4.1 Series connection:

In order to increase the voltage of the PV modules, we should connect the appropriate number of PV cells in series to get the right voltage. For this purpose, let us consider the single-diode model of the PV cell. As explained in the previous sections, I–V, the characteristics of a single-diode PV cell model are given below.

$$I = I_{SC} - I_s\left[\exp\left(\frac{V+I.R_S}{nV_t}\right)-1\right]-\frac{V+I.R_S}{R_{sh}} \tag{I.19}$$

In figure I.10 (b) is illustrated the method of connecting the circuit equivalent of a simple single-diode model to extract the value of $I_M\ and\ V_M$ when N_S identical series-connected PV cells operating under the same conditions.

$$I = I_{SC} - I_s\left[\exp\left(\frac{V_M+.N_S I_{SC}.R_S}{N_S.nV_t}\right)-1\right]-\frac{V_M+N_S.I.R_S}{N_S.R_{sh}} \tag{I.20}$$

$$V_M = N_S.V \tag{I.21}$$

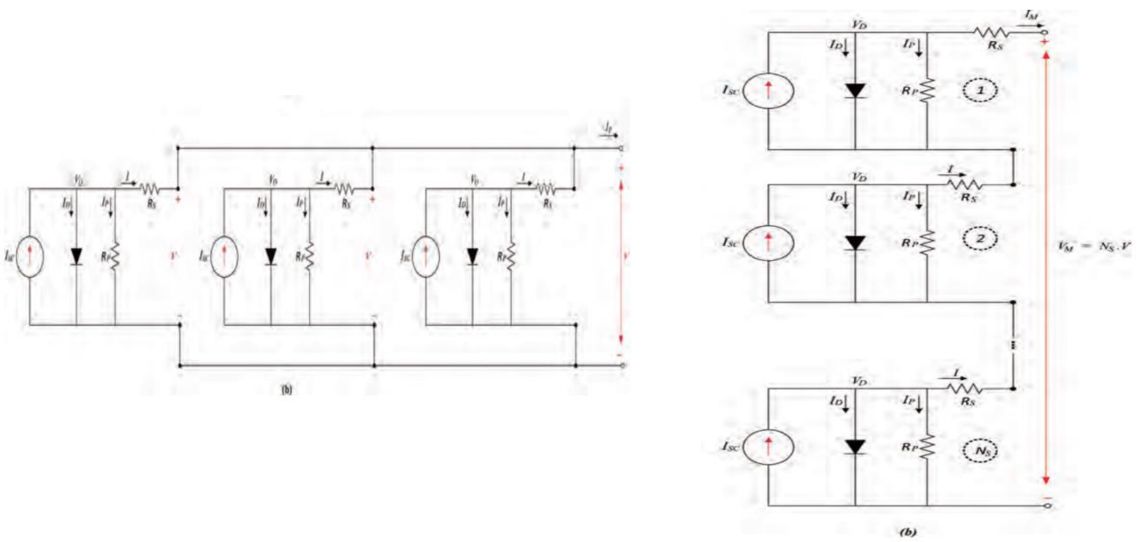


Figure I.10: Illustrated equivalent circuit in series (b) and circuit in parallel (a).

I.3.4.2 Parallel connection:

In order to increase the current of the PV modules, we should connect the appropriate number of PV cells in parallel to get the right current I_P,but the output voltage V_p will equal to the voltage of single PV cell. Again, for this purpose, let us consider the a single-diode model to of the PV cell. In the figure I.10 (A) is illustrated the method of connected the circuit equivalent of a simple single-diode model extract the value of $I_p\ and\ V_p$ when N_p identical parallel-connecting PV cells operating under the same conditions. I-V the characteristics of the single-diode PV cell model are given below as:

$$I_p = I_{SC}.N_p - N_p.I_s\left[\exp\left(\frac{V_M+N_p\backslash I_p.R_s}{nV_t}\right)-1\right]-\frac{V_M+N_p\backslash I_p.R_S}{N_p\backslash R_{sh}} \tag{I.22}$$

$$V_P = V_{cell} = V \tag{I.23}$$

I.3.4.3 Impact on voltage and current:

The influence of photovoltaic modules and arrays, whether connected in series or parallel, on voltage and current is significant, as the electrical behaviour of photovoltaic systems is heavily contingent upon the connection method employed. This table, I.2, demonstrated the effect on voltage and current.

Table I.2: Illustrated the impact of voltage and current in parallel and series connections.

Configuration	Voltage	Current	Power
Series	Increases $V_{module} = Ns \times V_{cell}$ $V_{array} = Ns \times V_{module}$	Constant $I_{module} = I_{cell}$ $I_{array} = I_{module}$	Increases $P_{module} = I_{cell} \times V_{module}$ $P_{array} = I_{module} \times V_{array}$
Parallel	Constant $V_{module} = V_{cell}$ $V_{array} = V_{module}$	Increases $I_{module} = Np \times I_{cell}$ $I_{array} = Np \times I_{module}$	Increases $P_{module} = I_{module} \times V_{cell}$ $P_{array} = I_{array} \times V_{module}$
Series-Parallel	Increases $V_{array} = Ns \times V_{module}$	Increases $I_{array} = Np \times I_{module}$	Maximum $P_{mpp} = I_{mpp} \times V_{mpp}$

I.4 The I–V and P–V Characteristics:

I.4.1 Key Electrical Parameters:

I.4.1.1 Short-circuit current (I_{SC}): is the value of the current electricity when the electrodes are short-circuited from the PV cell is illustrated in figure I.10. It occurs when the solar cell absorbs the sunlight falling on it and can be expressed in terms of value density from current (J_{SC}) multiplied by the area of solar cell. The *(I_{SC})* used to determine the performance and the efficiency of photovoltaic panels. This value is much higher than I_{mp}, which relates to the normal operating-circuited current and is expressed as follows [11]:

$$I_{SC} = J_{SC} \times A \text{ or} \tag{I.24}$$

Note that V=0.

$$I = I_0 \left[exp\left(\frac{qV}{nkT}\right) - 1 \right] - \mathrm{I_L} \tag{I.25}$$

$$\mathrm{I} = \mathrm{I_L} \tag{I.26}$$

Where:

I_0:is the reverse saturation current of the diode.

q: is the charge on the electron (1.602$\times$ 10^{-19} C).

k: is Boltzmann's constant (1.38$\times$ 10^{-23} m^2 kg/s^2 K).

T is the temperature in Kelvin.

The parameters I_0, I_L and n are dependent on the materials and design from solar cells.

I.4.1.2 Open-circuit voltage (V_{oc}):is the value of voltage measured between the electrodes of the PV cell is illustrated in figure I.11; it occurs when the solar cell absorbs the sunlight and there is no current flow. This value is much higher than V_{mp}, which relates to the operation. V_{OC} is the point at which there is an open circuit. As we can see, the voltage is as high as it can get, but the current is zero. Since voltage is multiplied by current, she is power zero, then is expressed as follows [11]:

V_{oc} × Zero amps = Zero power

$$V_{OC} = \frac{nkT}{q} ln\left(1 + \frac{I_{SC}}{I_0}\right) \tag{I.27}$$

Where: Kt/q represents the thermal voltage.

T: is the temperature.

n: is the ideality factor.

I_0 is the dark saturation current.

I_{SC} :is the light produced current.

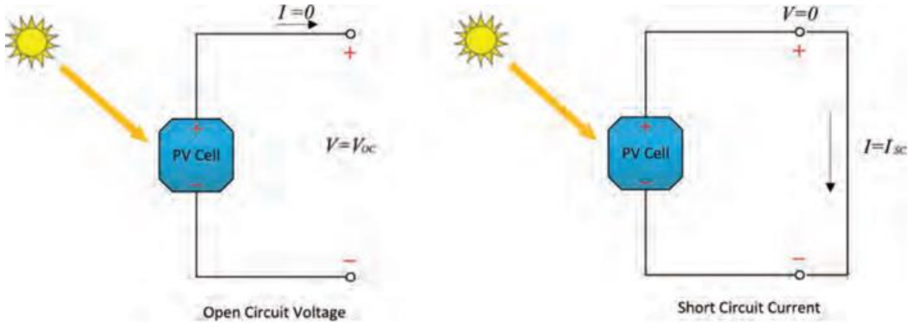


Figure I.11: The diagram illustrated the solar cell in the open-circuit voltage and short-circuit current.

I.4.1.3 Maximum Power Point (MPP):

Maximum power point (MPP or P_m, Pmax): It is the maximum power that a solar cell produces under the standard test condition (STC). The higher the Pm, the better the cell. It is given in terms of watt (W). Since it is maximum power or peak power, it is sometimes also referred to as Wpeak or Wp. A solar cell can operate at many current and voltage combinations. But a solar cell will produce maximum power only when operating at a certain current and voltage. This maximum power point is denoted in Figure I.12 as Pm. Normally, the maximum power point for an I-V curve of solar cells occurs at the 'knee' or 'bend' of the curve. In terms of expression, Pm is given as

P_m or $P_{max} = I_{mp}$ x V_{mp}, where I_m is current at max power point and V_m is voltage at max power point.

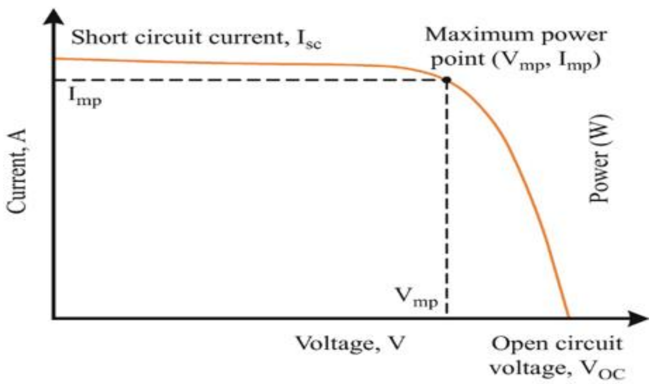


Figure I.12: The characteristic curve of a solar module illustrated the short-circuit current (I_{sc}), open-circuit voltage (V_{OC}), and maximum power point (MPP).

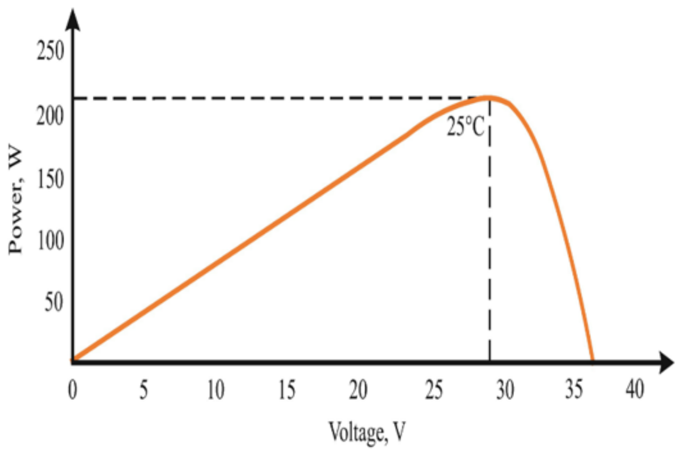


Figure I.13: The curve illustrates the characteristic P-V of the solar module.

I.4.1.4 Fill Factor (FF):

The Fill Factor (FF) is a vital parameter in the performance evaluation of the PV cell in serving as a key indicator of effectiveness and efficiency. The FF can be clearly defined as the proportion of the maximum power (Pmax) of a solar cell to the product of I_{sc} *and* V_{oc}. Graphically, the FF is a factor measure of the squareness level of the I-V curve of the PV cell. Namely, it is the ratio of the largest rectangle area under the IV curve to the largest rectangle ($V_{oc}\times I_{sc}$)which covers the I-V curve. Basically, the greater the FF, the performance of the PV cell is better. The FF is graphically illustrated, and it is clear from that the FF can be expressed mathematically in Figure I.14. [8]

The expressed mathematically as

$$FF = \frac{P_{max}}{V_{oc}.I_{sc}} = \frac{\eta \,.\, A_c.G}{V_{OC}.I_{sc}} = \frac{V_{mp}.I_{mp}}{V_{oc}.I_{SC}} \tag{I.28}$$

Where:

P_{max}: maximum power.

V_{oc} : Open-circuit voltage.

I_{sc}: short-circuit current.

V_{mp} : maximum power voltage.

I_{mp}: maximum power current

η: the cell efficiency

A_c: cell area (m²)

G: the solar radiation (W/m²).

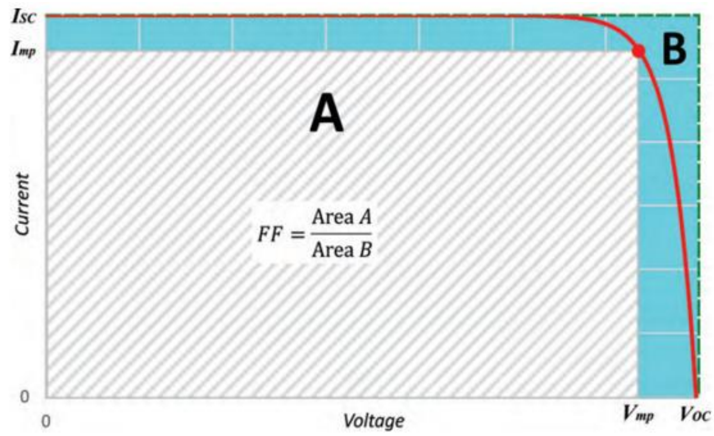


Figure I.14: The illustration of FF on the I-V curve

I.4.1.5 Efficiency (η):

The PV cell efficiency is a major factor in understanding the performance and effectiveness of PV cell. It can be defined as the ratio of output maximum power to the input power due to incoming solar radiation on a PV cell surface. Cell efficiency is not only dependent upon solar radiation but also depends on the solar spectrum, the intensity of incident sunlight, and PV cell temperature. it is vital to control the conditions under which the efficiency is measured to compare the PV cell systems' performance one to another. For example, terrestrial photovoltaic cells are measured under air mass (AM1.5) conditions at a temperature (25°C), while PV cells for space or satellite usage are measured under AM0 conditions at another temperature.

The PV cell efficiency (**η**) is typically expressed as a percentage, and then we can define it mathematically as follows:

$$\eta = \frac{Maximum\ Electric\ Power}{inpute\ power} = \frac{P_{max}}{Area\times\ Irradiance} = \frac{V_{OC}.I_{SC}.FF}{A_C.G} \tag{I.29}$$

η: cell efficiency, P_{max}: maximum electrical power, A_c represent cell surface area, G represents the incoming solar radiation, V_{OC} is the open-circuit voltage, I_{SC}represents the short-circuit current, FF is the Fill Factor.

Some other important factors such as solar radiation, temperature, and wind speed, which are responsible for the performance and efficiency of a solar cell, are comprehensively covered in the next section. [8]

I.4.2 Effect of Irradiance on Characteristics:

the characteristic curves for the solar cell in Figure 1.14, with irradiance as a parameter. Here, the short-circuit current I_{SC} is proportional to irradiance, whereas open-circuit voltage increases only slightly as irradiance rises. This also means that solar cell voltage can be quite high even in the presence of very low irradiance – for example, at dusk. This fact should be taken into account when installing and servicing PV systems that exhibit relatively high voltages.

Inasmuch as the ratio between voltage at the MPP $V_{MPP}$and open-circuit voltage V_{OC} fluctuates only slightly, the MPP voltage is also somewhat lower for lower irradiance. In addition, under extremely low irradiance conditions, the voltage flowing through the parallel resistance Rp weighs more heavily on balance and results in an additional voltage reduction at the solar cell. On account of this reduced V_{MPP} in the presence of low irradiance and the current flowing through Rp, solar cell efficiency is also lower in such situations. Figure15. I.

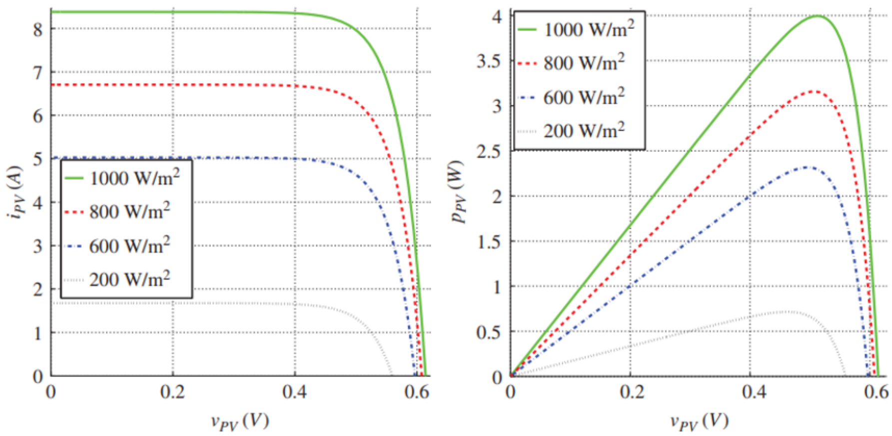


Figure I.15: Characteristic curves I =f(V) for the solar cell, with cell temperature of 25°C and irradiance G as a parameter (AM1.5 spectrum).

I.4.3 Effect of Temperature on Characteristics:

the characteristic curves for the solar cell in Figure I.16., with cell temperature as a parameter. In Si solar cells, open-circuit voltage decreases as temperature increases, by around 2 to 2.4 mV/K (the temperature coefficient of V_{OC} is +0.3% per K up to +0.4% per K); this decrease is attributable to the fact that diode threshold voltage in the equivalent circuit diagram decreases as for a standard Si diode. In this context, MPP voltage of course decreases as temperature rises. The fill factor FF likewise decreases as temperature rises, the FF temperature coefficient is roughly +0.15% per K). Inasmuch as short-circuit current increases only slightly as temperature rises (the I_{SC} temperature coefficient is only about þ0.04 to þ0.05% per K), power Pmax at the MPP also decreases as temperature increases. For crystalline silicon solar cells, the temperature coefficient T for Pmax ranges from about +0.4 to+ 0.5% per K. Inasmuch as PV system cell temperature can be 20 to 40°C above the ambient temperature in the presence of high irradiance, this can provoke a considerable power drop and thus reduced efficiency at high temperatures; these parameters are often underestimated.

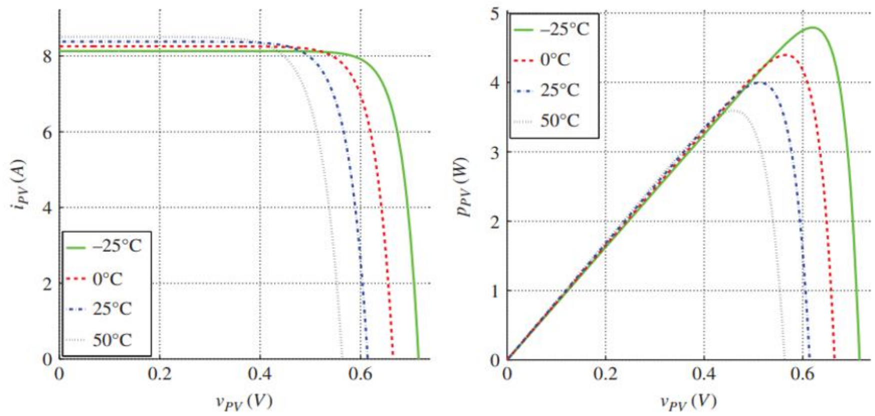


Figure I.16: Characteristic curves I = f(V) for the solar cell as in Figure I.15, with irradiance G=1 kW/m² (AM1.5 spectrum) and with cell temperature as parameter (variable).

I.5 Classification of PV Faults:

Photovoltaic (PV) faults are generally classified into electrical faults, environmental faults, and degradation-related faults.

I.5.1 Environmental Faults:

I.5.1.1 Partial shading:

This condition in which PV panels connected in PV array are subjected to not the same irradiance level is called as partial shading. So, the partial shading in any solar cell or any string of cells can be a major disadvantage in the solar cell. Or in other word the PV modules are very sensitive to shading, hence, the modules that which are exposed to lesser solar irradiance tend to generate low level of current compared to PV panels which are fully irradiated and hence this leads to the problem of current imbalance between the PV panels connected in the same series connection. The partially shaded PV panels are compelled to carry higher amount of current which causes them to get reverse biased and start acting as load to the existing system. The overall power output decreases and also causes increase in the temperature of the partially shaded panels. The rise in temperature leaves Hot-Spot which damages the PV panels and effect in their the MPP and the efficiency.

I.5.1.2 Dust accumulation:

The accumulation of dust on solar photovoltaic modules results from atmospheric deposition onto the module's surface illustrated figure I.17. This process is significantly influenced by wind speed and the inherent characteristics of the dust sources, particularly when these sources are in close proximity to the PV module. Dust accumulation occurs through two main mechanisms, dry accumulation, in which dust particles adhere to the solar panels without moisture, and wet accumulation, where humidity from rain, fog, or snow enhances dust adhesion. This accumulation of dust adversely affects the performance of the PV system, leading to reduced efficiency.

Figure I.17: dust accumulation of the PV module.

I.5.1.3 Snow and humidity:

The humidity and snow are among the factors that affect solar panels, especially the materials used in manufacturing the PV panels. Places where snow falls frequently have a high humidity level, which weakens the internal and external components of the photovoltaic panel unit, like the back-sheet foil and the encapsulation layer up to the intercell connectors and the cell metallization. All those metallic components can subsequently undergo corrosion reactions. it may lead to weakening the performance and efficiency of the solar cell, in other words it is the deterioration of the PV panels.

I.5.2 Electrical Faults:

I.5.2.1 Series resistance increase:

The first, the series resistance (Rs) they represent the sum of resistance due to all the components that path of current. This includes the emitter, base, resistance of metal contact and semiconductor-metal contact resistance. It is desirable to have the value of series resistance as low as possible. So, the increased of the value of series resistance (Rs) as result of cell solar degradation an example to the corrosion and aging the PV cell or environmental faults that as impact in reduce in the short-circuit current Isc region. The open-circuit voltage (Voc) remains relatively unaffected, but the fill factor (FF) and maximum power point degrade substantially.

I.5.2.2 Shunt resistance decrease:

Againe the shunt resistance (Rsh) is due to the leakage across the p-n junction. It could be due to a shunt around the periphery of the cell or due to the crustal defect or precipitates of impurities in the junction region. It is desirable to have the value of shunt resistance as high as possible. a parallel leakage path, resulting in increased current loss particularly at low voltages. This manifests as a steeper slope in the low-voltage region of the I-V curve and directly reduces both Voc and FF.

I.5.2.3Bypass diode failure:

PV modules include bypass diodes to protect the cells from hot spots if the module is partly shaded, either from soiling or a physical obstacle see figure I.18. Hence, bypass diodes in a module allow the non-shaded cells to continue power generation while protecting the shaded cells from being overheated and damaged. Bypass diodes can fail in either short-circuit or open-circuit mode. If it fails short circuited, it will pass current whenever the sun is shining so the cells it is protecting will no longer produce any useful electricity. If the diode fails open circuit, it acts as though it were not in the electric circuit at all and therefore provides no reverse bias hot spot protection to the module. It is fairly easy to see when a diode short circuits as the module no longer produces the rated amount of electricity and the diode is on and hot whenever the sun shines. On the other hand, an open-circuited diode is difficult to observe since it has no impact on an operating module. To determine whether bypass diodes are operating or not requires a module-by-module shading test that is very time-consuming.

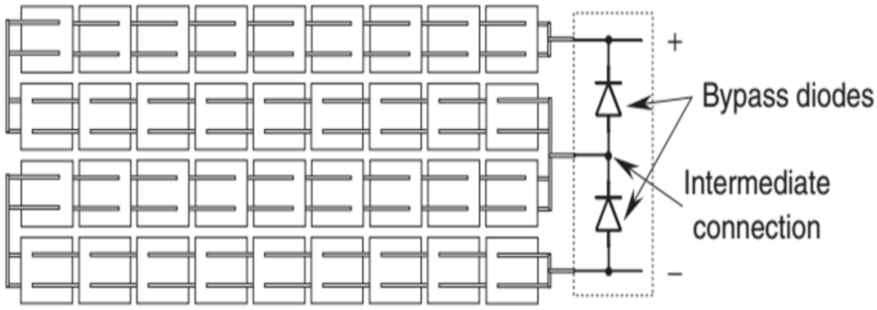


Figure I.18: This diagram illustrates two bypasses connected into 36 cells in series.

I.5.2.4 Open-circuit faults:

Open-circuit faults result from broken interconnections, failed solder bonds, or connector failures. The entire string carrying the open circuit goes to zero current, causing complete loss of contribution from that branch. In a parallel configuration, this reduces the total array current proportionally.

I.5.2.5 Short-circuit faults:

Short-circuit conditions occur between string conductors or between a string and the grounding conductor. They produce abnormal current flow, bypassing normal circuit paths, and can pose significant fire and equipment damage risks if not promptly isolated.

I.5.3 Structural and Mechanical Faults:

I.5.3.1 Cell cracks:

The solar cell, which has a thickness of a few microns, is prone to cracks due to the fragile structure. These cracks are a common defect in solar panels. There are results either during manufacturing or due to exposure to mechanical force or thermo-mechanical stress, as illustrated in figure I.19. These cracks may be visible to the naked eye or even need special techniques to locate the cracked cell in a module. During field operating conditions, the cracks act as favourable sites for moisture and air ingress, resulting in metallisation corrosion and the degradation of encapsulants due to moisture ingress along the cracked regions.

Figure I.19: illustrated the crack of the solar cell.

I.5.3.2 Delamination:

Delamination in the PV solar modules is the separation of different layers within the module package. It is one of the structural and mechanical fault types encapsulant degradation. This is due to the climatic conditions surrounding the solar panels, like UV radiation, high temperature, humidity, thermal cycling and oxygen exposure. This weakens the bond between the solar cell and the packaging material, resulting in a gap between the two layers.

Delamination can occur due to bad manufacturing. like Poor lamination process and adhesion resulting in the insufficient flow of Ethylene Vinyl Acetate (EVA)under interconnect ribbon tabbing could leave voids between the layers. Also, the chemical reactions involving the degradation of different components of the module release gases, which get trapped between the cell and the EVA layer. The trapped gases under high-temperature conditions could expand, resulting in the widening of the delaminated region. Moreover, the additives added in the EVA to strengthen the adhesive bonds to protect the cell from dust, humidity, Water vapor, cracks, external shocks but in to the degrade under hydrothermal conditions. When these additives deplete, the interface becomes prone to degradation under external environment conditions.it effect to the optical losses Decrease in short-circuit current Isc reduction in maximum power Pmax and lower efficiency, local overheating, See Figure I.20.

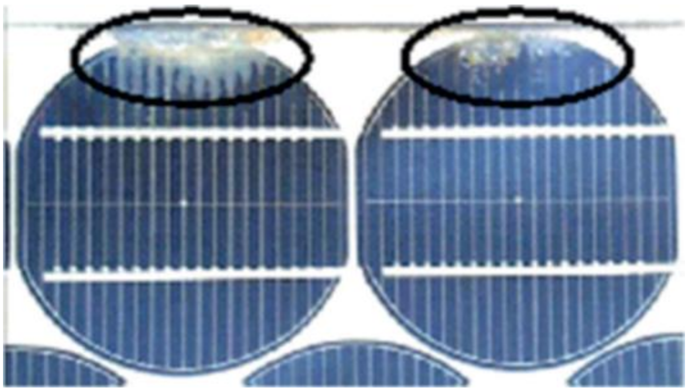


Figure I.20: delamination solar cell module

I.5.3.3 Solder bond failure:

Failed solder bonds at the cell busbars or ribbon interconnects increase the local series resistance of the affected cells see figure I.21. Progressive bond degradation manifests as a gradual fill factor reduction and can eventually lead to open-circuit conditions. These failures are often thermally induced through repetitive diurnal temperature cycling.

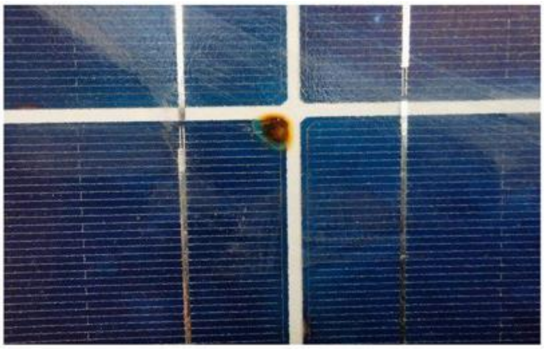


Figure I.21: The solder bond failure in PV module.

I.6 Detailed Analysis of Major PV Fault Mechanisms:

PV faults mainly include partial shading, open-circuit faults, short-circuit faults, hotspots, and module degradation. These faults alter the electrical behavior of PV modules by reducing voltage, current, and power output. If not detected early, they can decrease system efficiency, accelerate aging, and compromise operational reliability.

I.6.1 Partial Shading Effect:

I.6.1.1 Hot spot formation:

Hot-spots in PV modules represents a broad defect type, with many presentations and underlying causes, here we'll consider three broad categories that cause hot-spots. This category order also roughly follows an increasing effort of investigation and remediation, and therefore source of concern for owners and operators. See the figure I.22.

1. **Shading/soiling:** overhead objects (ex. trees, poles, etc.), vegetation overgrowth, surface fouling, foreign objects on surface.
2. **Mechanical damage:** broken glass, broken/bent frame, collisions of modules with each other or other objects, improper fixturing.
3. **Internal module failures:** cell material defects (ex. shunts, high series resistance, etc.), cell cracks, local de-lamination, poor solder joints.

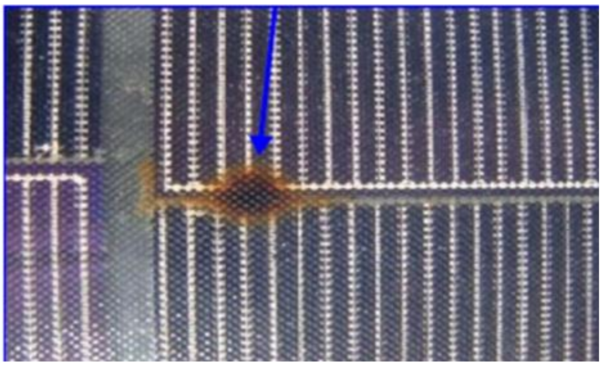


Figure I.22: Hot spot failure in a solar cell.

I.6.1.2 Multiple MPP phenomenon:

The phenomena of the multiple maximum power pointe caused by the mismatch of PV modules rise when all are connected in series. In one case, two power peaks are noticeable in the. Neither can represent the true available power, which is equal to the sum of the individual maximum powers of the two PV modules. The mismatch is due to a bottleneck effect caused by the under-performing PV panel. However, the case study is based on a simple shading scenario with only two PV modules in series connection. In reality, the mismatch condition can be very complicated when more than two panels are connected in series in each string. In centralized PV systems, the strings are in parallel connection to form an array. This might result in multiple peaks and complicated scenarios that are difficult to identify and predict, in order to minimize the effect of mismatch, conditions are independent MPP at string, panel, and even submodule levels. The approach is commonly referred to as Distributed MPPT (DMPPT). this system creates an independent channel for each power contributor, with power collected at the final stage for grid injection or load consumption. The mismatch effect is ultimately eliminated or reduced according the level at which MPPT is applied.

I.6.1.3 Energy mismatch:

The current mismatch between shaded and unshaded modules in a series string forces the string current to the level of the weakest (most shaded) module. The power deficit from energy mismatch can be substantially larger than that attributable to the irradiance loss on the shaded cells alone. This multiplicative effect makes partial shading one of the most economically significant fault types in large PV installations.

I.6.2 Series Resistance Variation:

I.6.2.1 Causes (aging, corrosion):

The corrosion and aging have impact on the series resistance variation in the PV cell, the series resistance has represented all the contact and semiconductors resistance like fingers, the busbars or the metallized components of PV cell because the degradation factors as aging and corrosion can be influenced which to the series resistance increased this increase is undesirable it directly affects in the current of PV module and turn on the power and fill factor (FF).

I.6.2.2 Effect on fill factor:

The impact of the series resistance in fill factor (FF). Yes, the resistance has a massive effect on the value of the FF ranges from 50–82%. For a Silicon PV cell, the FF is usually about 80%. The value increases with a higher shunt resistance and a lower series resistance. A higher value of FF is desired because it is a direct indicator of the quality of the PV cell.

I.6.2.3 Mathematical interpretation:

An equation for the FF as a function of series resistance can be determined by noting that for moderate values of series resistance, the maximum power may be approximated as the power in the absence of series resistance minus the power lost in the series resistance. The equation for the maximum power from a solar cell then becomes:

$$P'_{MP} \approx V_{MP}I_{MP} - I_{MP}^2 R_S \tag{I.30}$$

$$P'_{MP} = P_{MP}\left(1 - \frac{I_{SC}}{V_{OC}} R_S\right) \tag{I.31}$$

$$P'_{MP} = P_{MP}\left(1 - \frac{R_S}{R_{CH}}\right) \tag{I.32}$$

defining a normalized series resistance (r_S) as:

$$r_S = \frac{R_S}{R_{CH}} \tag{I.33}$$

R_S is series resistance in the, R_{CH} is the chrematistics series resistance.

Assuming that the open-circuit voltage and short-circuit current are not affected by the series resistance allows the impact of series resistance on FF to be determined:

$$P'_{MP} = P_{MP}(1 - r_s) \tag{I.34}$$

$$V'_{OC}I'_{SC}FF' = V_{OC}I_{SC}FF(1 - r_s) \tag{I.35}$$

$$FF' = FF(1 - r_s) \tag{I.36}$$

The above equation the fill factor which is not affected by series resistance is denoted by FF_0 and FF' is called FF_S. The equation then becomes:

$$FFs = FF_0\,(1 - r_S) \tag{I.37}$$

FFs is fill factor of with series resistance increased, FF_0 is fill factor ideal in the (STC).

I.6.3 Shunt Resistance Variation:

I.6.3.1 Leakage current behavior:

After connected a large number of the photovoltaic (PV) solar cells modules. In a PV system, in the series and parallel arrays into different strings. This arrangement of modules produced the high voltage in this case the modules must de grounded for safety reasons. A current is generated under this voltage stress, is cells a leakage current. this leakage current effect to the availability of an adequate number of ions (i.e., Na+) on the solar cell surface leads to potential induced degradation (PID). This results in the degradation in the performance of a solar cell and to the different paths available for leakage current to flow. This leakage current depends on many factors, which can be categorized as module components and environmental conditions. Temperature and humidity are important factors in deciding the amount of leakage current.

I.6.3.2 Impact on low-voltage region:

the shunt resistance R_{sh}has no effect on the short-circuit current, but reduces the open-circuit voltage. Conversely, the series resistance RS has no effect on the open-circuit voltage, but reduces the short-circuit current. Sources of series resistance include the metal contacts, particularly the front grid, and the front grid.

I.6.4 Bypass Diode Malfunction:

the bypass diodes are included in PV module to protect them from the partial shaded, because the PV cells in real-life conditions, shading may be caused by a nearby object, such as a tree, adjacent building, chimney, soiling or a physical obstacle. We know, in case series connection of the solar cells their current value is total cell current in the partially shaded of one cell this means that the current generated in the shaded cell will decrease significantly. Therefore, this cell determines the maximum current flowing through the unit, because, the shaded cell does not generate energy but begins to dissipate power and heats. These problems caused by partial shading can be prevented by including a bypass diode in module, in doing a diode blocks current when it is under a negative voltage, bit allows current to pass when it is a positive voltage.

I.6.5 Thermal stress consequences:

The consequence of the thermal stress from the bypass diode malfunction, which are essential to protecting the module during partially shaded conditions. If bypass diodes fail in open circuit, partial shading can cause severe hotspots and fire hazards. modules, failure of one bypass diode in short-circuit results in approximately one-third reduction of output power, certainly a module failure. Even with working bypass diodes, continuous operation of cells in reverse bias means the diodes will operate at high temperatures over long time periods which can cause permanent degradation of the diodes are working toward developing enhanced understanding of the stressors and field-failure mechanisms in commonly used silicon-based Schottky bypass diodes. Knowledge of long term, shading-induced degradation mechanisms in cells and bypass diodes in the field is essential for estimating the service life of PV modules.

I.7 Impact of Faults on PV System Performance:

I.7.1 Power loss estimation:

Photovoltaic (PV) systems operate under varying environmental and electrical conditions for many years. During operation, different faults and degradation mechanisms appear and progressively affect system performance. The fault of PV panel or failure, that are driven in power losses, like, a bypass diode failure, open-circuit fault and short-circuit fault.

I.7.2 Efficiency degradation:

The efficiency of the PV modules or system is one of the components expressed in terms the performance of the PV cell in standard reporting conditions (SRC) or standard test conditions (STC).

As it is known the PV conversion efficiency (η) is calculated from the measured maximum or peak PV power (Pmax), device area (A), and total incident irradiance (G):

$$\eta = \frac{Maximum\ Electric\ Power}{inpute\ power} = \frac{P_{max}}{Area\times\ Irradiance} = \frac{V_{OC}.I_{SC}.FF}{A_C.G} \quad (I.38)$$

But there are always factors or the failures that reduce the efficiency of solar cell another since the efficiency degradation, like to the failures related to the output current , output voltage and maximum power, there have also the failures related to electrical degradation (increased series resistance, reduce shunt resistance and bypass diode failure), the thermal degradation, the mechanical degradation and environment faults (particle shading, dust,…) and there's the temperature influenced in solar cell efficiency approximately when to flowing:

$$\eta(T) = \eta ref[1 - \beta(T - Tref)] \quad (I.39)$$

β temperature coefficient,T operating temperature, T_{ref}= reference temperature. Finally, the consequence of the efficiency degradation is reduced energy generation lower return on investment, faster system aging and reduced operational lifetime.

I.7.3 Reliability concerns:

The reliability of photovoltaic panels is important in PV system, the question of the lifespan of the PV cell module when exposed to external operating remains a challenge because the faults or the failures and external stress factors. The most accurate way assesses the reliability of PV is to monitor their performance in actual operation.

I.7.4 Economic impact:

The economic impact in terms of fault on PV system performance is important because always, when we say PV say electricity production, because the photovoltaic panels are the basic components of any station electrical production by PV panels, from small rooftop systems with a capacity of a few kilowatts to G.Watte plants that require millions of units. Recent studies that cost of the productid one the kilowatt-hour(kwh) between 2and 6 cents euros from the photovoltaic electricity, depending on the size and location of the power station. and from it the impact of fault and degradation on PV system and diagnosis not perfect, has caused decreased the production and the cost of the diagnostic and maintenance.

I.8 Diagnostic Challenges in PV Systems

I.8.1 Similar electrical signatures:

One of the most prominent issues is that PV systems can produce similar electrical signatures even under different fault conditions. This phenomenon complicates the diagnostic process, as it can lead to confusion and misinterpretation of data. Without precise identification, technicians may misdiagnose problems, resulting in improper maintenance actions that could exacerbate the situation.

I.8.2 Environmental variability:

PV systems are also greatly affected by external environmental elements such temperature variations, shade caused by other buildings or trees, and different weather conditions. Such changing circumstances may greatly affect the amount of electricity a PV system produces, masking the presence of hidden problems that need to be fixed. e.g. a rapid drop in the amount of solar light may lead to a reduction in energy output which might be misinterpreted as a failure when it is just a natural environmental effect, also mentions that dynamic environmental variables provide challenges in defect identification, and hence, the algorithms should be resilient to fluctuating irradiance and temperature. In reality, this means a model must distinguish actual defects from regular weather-driven volatility.

I.8.3 Data imbalance:

A critical challenge in fault diagnosis is the often-imbalanced nature of the data collected for analysis. Certain types of faults may be significantly underrepresented in the datasets, which can limit the effectiveness of diagnostic tools and machine learning models. This imbalance hampers the system's ability to adequately learn from and recognize less common fault conditions, ultimately affecting decision-making processes or determine the type of the fault.

I.8.4 Real-time monitoring limitations:

For rapid detection of faults real-time monitoring systems are necessary but often they have limitations in processing capabilities especially with large data. Failure to analyze this data efficiently can delay fault detection and response mechanisms. Such delays may also result in long downtimes, higher repair costs, and damage to the PV system.

To address these diagnostic problems effectively, sophisticated methods such as artificial intelligence and machine learning are essential. These technologies improve the detection and diagnosis capabilities and make them more accurate. Furthermore, the development of broad data collection and analysis schemes can mitigate the problems caused by data imbalance and environmental variability, thus enhancing failure resilience and performance of PV systems.

Chapter II:
Modeling and Simulation of PV Panels Under Fault Conditions

II.1 Introduction:

The reliability and sustainability of photovoltaic (PV) systems represent some of the most critical challenges in the renewable energy sector, as faults and non-uniform shading conditions lead to a significant decline in production efficiency and an increase in operational maintenance expenses.[25]Consequently, simulation-based fault diagnosis emerges as an imperative necessity, not only to minimise system downtime but also to provide a safe and cost-effective testing environment. That allows for the study of complex fault scenarios that are difficult to replicate in the field [25]. To achieve this high degree of model precision, MATLAB/Simulink serves as a benchmark tool due to its specialised libraries and their capacity to simulate the non-linear physical behavior of solar cells under diverse environmental conditions [23,26]. The simulation framework in this thesis adopts a systematic approach that begins with constructing a healthy reference model of the PV panel based on the single-diode model equation. [23] This is followed by the software-based injection of specific faults, such as partial shading or electrical degradation, and monitoring their impact on the current-voltage (I-V) and power-voltage (P-V) characteristic curves. [24] Ultimately, this analytical path aims to extract the distinct signatures and "fingerprints" of each fault type, paving the way for the development of monitoring and diagnostic strategies that ensure the long-term stability of the energy system.

II.2 PV panel modelling in MATLAB/Simulink:

I.2.1 Normal Operation Model:

The electrical behavior of a photovoltaic (PV) cell is governed by the standard diode equation, which relates the output current (I) to the terminal voltage (V). According to Duffie and Beckman [31], the current generated by the PV cell can be mathematically expressed as:

$$I = I_{ph} - I_0\left[\exp\left(\frac{q(V+IRS)}{nkT}\right) - 1\right] - \frac{v+IRS}{Rsh} \tag{II.1}$$

Where:

- I_{ph}: Photo-generated current (A).
- I_0: Diode reverse saturation current (A).
- q: Electron charge ($1.602 * 10^{-19} c$).

- K: Boltzmann constant ($1.381 * 10^{-23} J/K$), which serves as physio relating the average relative kinetic energy of particles in a gas with the thermodynamic temperature of the gas [1].
- T: absolute temperature of the p-n junction (K).
- n: Diode ideality factor.
- R_s, R_{sh}: series and shunt resistances (Ω).

Based on the mathematical model described above, a computational representation was developed using the Sims cape Specialised Power Systems library in MATLAB/Simulink. The following block diagram (Figure 1) illustrates the implementation of the PV array and measurement blocks used to capture the electrical characteristics.

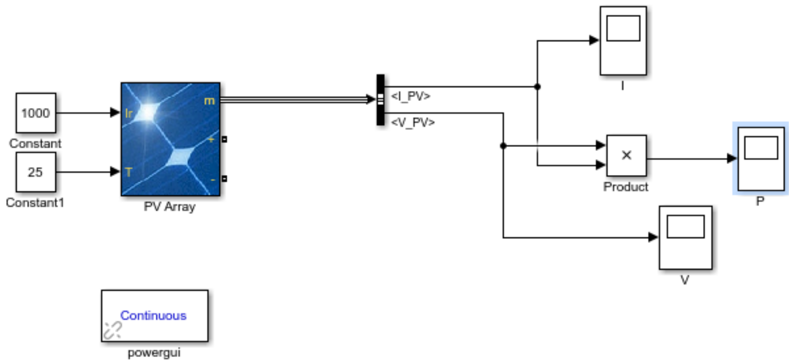


Figure II.1: Simulink block diagram of photovoltaic system module.

In the mathematical model of the PV cell, the Boltzmann constant (k) plays a fundamental role in linking the thermal energy of the p-n junction to the electrical output. The product of (k*T) determines the thermal voltage, which signficantly influences the exponential behaviour of the diode equation. This explains the physical phenomenon where an increase in temperature leads to a reduction in the open-circuit voltage (V_{OC}), thereby affecting the overall efficiency of the PV model. For this simulation, the values of the Boltzmann constant and the electron charge (q) were integrated to ensure the model's accuracy under varying climatic conditions [23, 24].

Table of Specifications of PV Array **:** illustrated the charactirstic PV panelin table II.1

Table II.1: charactirictic PV panel.

Parameter	**Symbol**	**Value**	**Unit**
Maximum Power	P_{max}	305.2	W
Open-circuit voltage	V_{OC}	64.2	V
Short-circuit current	I_{SC}	5.96	A
Number of cells	N_{cell}	96	-
Reference Temperature	T_{ref}	25	°C
Reference Irradiance	G_{ref}	1000	W/m^2

Fundamental physical constants used in the model in table II.2

Table II.2: fundamental physical of PV cell.

Unit	**Value**	**Symbol**	**Constant**
J/K	$1.3806*10^{-23}$	k	Boltzmann
C	$1.6022*10^{-19}$	q	Electron charge

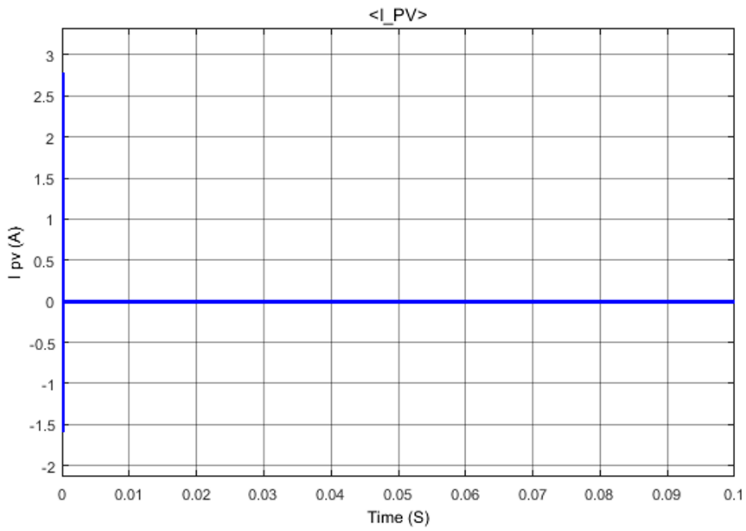


Figure II.2: PV array output current response under step-changing solar irradiance.

Figure II.2: illustrates the output current (Ipv) of the PV array during the simulation period. The simulation was conducted under a variable irradiance profile to test the dynamic response of the model. It is observed that the current magnitude changes instantaneously with the irradiance levels (from 1000W/m^2 down to 200W/m^2).

From a physical perspective, this linear relationship confirms that the photo-generated current is a direct function of the incident photon flux. The stability of the current at each step demonstrates that the model, which incorporates the Boltzmann constant (k) and the electron charge (q), correctly calculates the carrier generation rate with no oscillations during the transitions, which proves the numerical stability of the developed Simulink model.

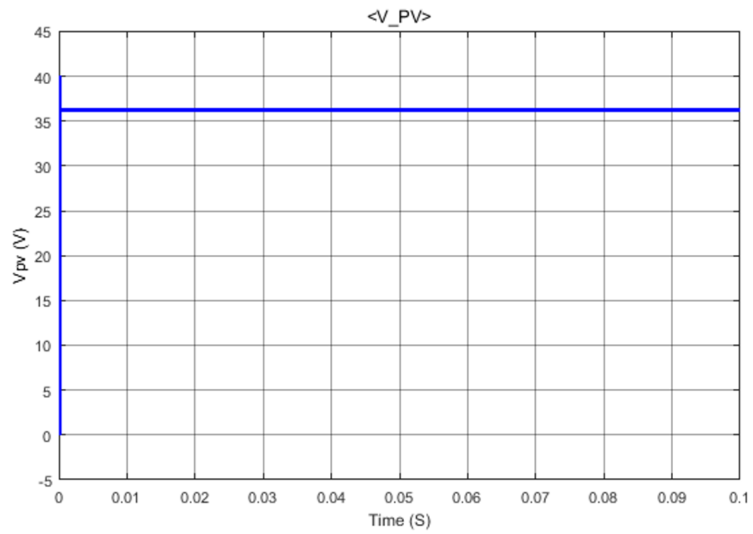


Figure II.3: PV array output voltage response under step-changing solar irradiance.

The terminal voltage profile (V_{pu}) obtained from the simulation is depicted in Figure II.3. Unlike the current response, the voltage demonstrates a remarkable resilience to irradiance fluctuation, maintaining a steady level of approximately **480 V.** The mathematical model established by [1] provides the following information. This behavoir is attributed to the logarithmic dependency of the open-circuit voltage on the solar intensity, which explains why the voltage does not drop as drastically as the current when the irradiance decreases. The slight transients observed at the transition points (e.g., at t=1.0s) are primarily due to internal capacitance and the non-linear nature of the diode equation. These results validate the steady-state accuracy of the model, ensuring a reliable DC input for the power conditioning stage.

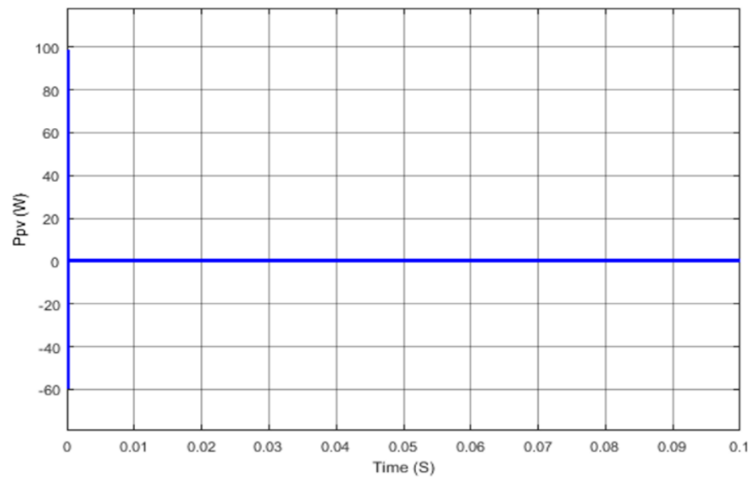


Figure II.4: Output power (P_{pv}) characteristics of the PV array under variable irradiance.

The output power characteristics (Ppv) shown in figure II.4 validate the successful integration of the PV array components. The power effectively tracks the irradiance steps, reaching its peak at the initial standard conditions[25]. This rapid and stable transition between power levels without sustained oscillations is a key indicator of a well-parameterised solar cell model. These results ensure that the model is ready for the next stage of the study, which involves the integration of an MPPT controller to optimise energy extraction.

II.2.2 PV Array Modelling:

II.2.2.1 Series configuration: A photovoltaic (PV) array is composed of multiple modules interconnected to meet specific power requirements. Series configuration is the primary method used to increase the system voltage, where the positive terminal of one module is connected to the negative terminal of the next. According to [23]. This arrangement ensures that the same current flows through all series-connected units while the individual voltages are summed.

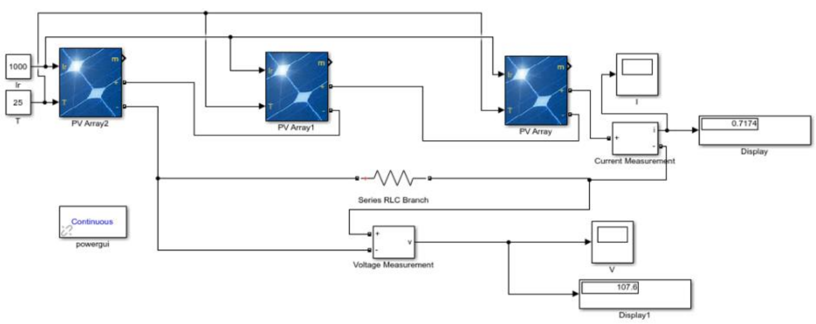


Figure II.5: Series-connected configuration.

Figure II.5 illustrates the electrical schematic of a PV array with modules connected in series. The diagram demonstrates the cumulative voltage gain across the string terminals while maintaining a uniform current flow through all interconnected modules.

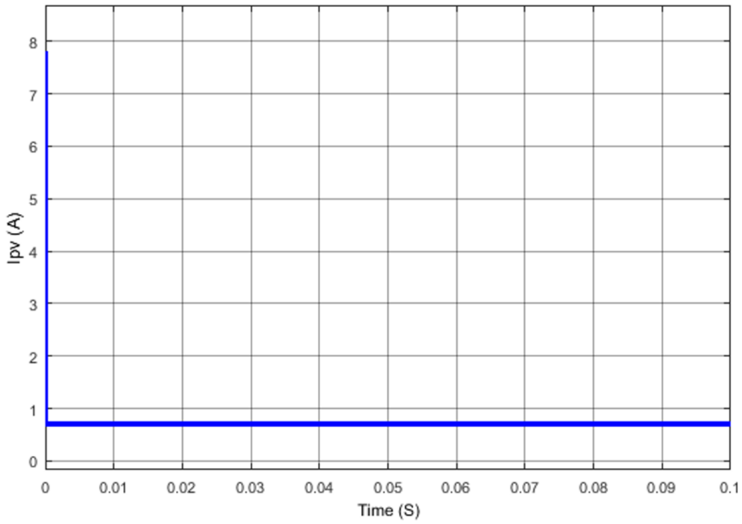


Figure II.6: PV array output current response under step-changing solar irradiance (3 modules in series connection).

As illustrated in the simulation results (figure above), the array total current (I_{up}) remains constant and stable at **0.7174 A** throughout the simulation period. In a series circuit, according to Kirchhoff's Current Law, there is only one path for charge flwo; thus, the current passing through the first module is identical to that passing through all interconnected modules.

- **Result:** the total system current $I\text{array} = I\text{module} = 0.7174A$
- **Explanation:** This explains the horizontal steady line observed in graph, indicating stable electron flow within the string [27].

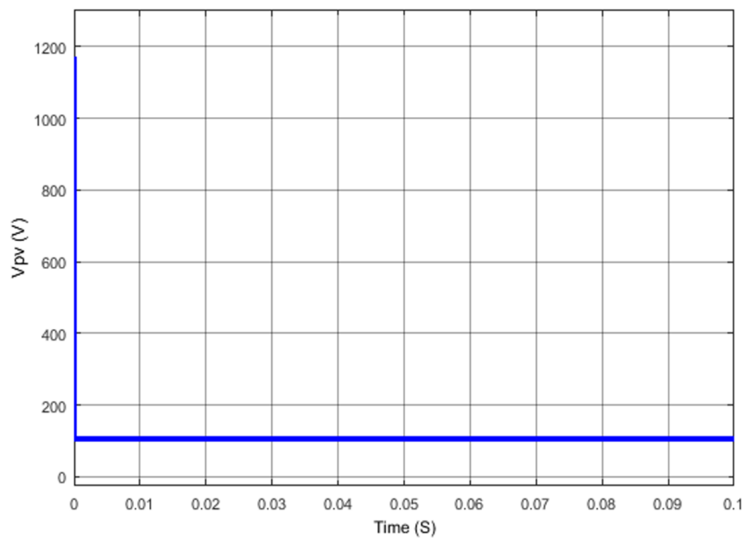


Figure II.7**:** PV array output voltage response under step-changing solar irradiance (3 modules in series).

Unlike the current, the total output voltage (V_{pu}) in a series circuit follows the principle of accumulation, as presented in Figure 7.II. Since each module functions as an independent electromotive power source, the total steady-state voltage monitored at the string terminals is the algebraic sum of the individual module voltages.

- **Result:** $V_{\text{total}} = V_1 + V_2 + V_3 = 107.6$ V.
- **Scientific Integrity:** [1]. States that this configuration is specifically designed to minimise electrical transmission losses; increasing the voltage allows the power to be transmitted with a lower current, thereby reducing cable heat dissipation ($P_{\text{loss}} = I^2$ R).

I.2.2.3 Parallel Configuration:

Parallel configuration is employed when the primary objective is to increase the system's total current while maintaining a constant voltage. This is achieved by connecting all positive terminals together and all negative terminals to a common junction point. [34], this configuration is essential for ensuring system compatibility with charge controllers or inverters that operate at low input voltage and high current levels.

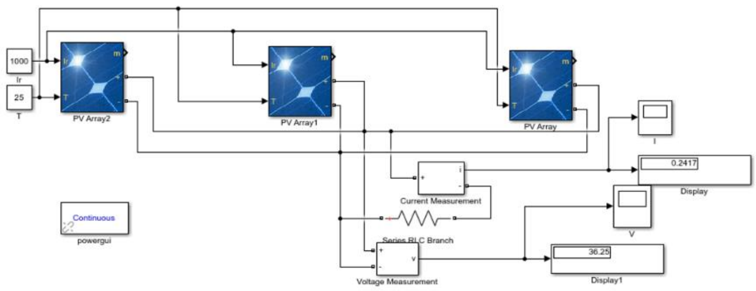


Figure II.8: Parallel Configuration.

The schematic should illustrate modules aligned side-by-side, with current lines from each module (I_1, I_2, I_3) feeding into a main busbar representing the total current ($I_{tota} = \Sigma I_n$). Meanwhile, the voltage is measured across the common junction points, ensuring that $V_{array} = V_{module}$.

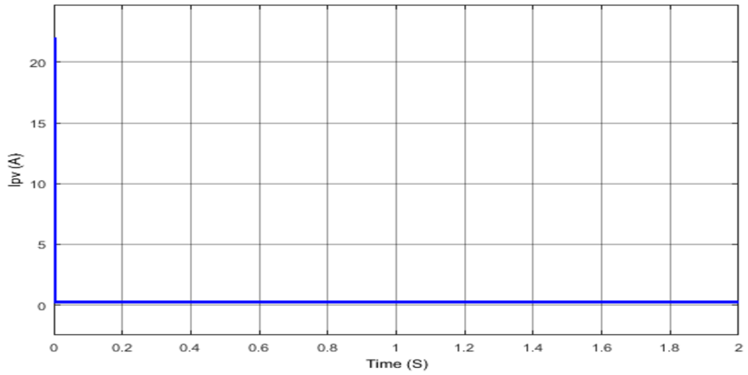


Figure II.9: PV array output current response under parallel configuration (3 single modules).

Figure II.9: illustrates the current response of a PV array with modules connected in a parallel configuration. It is observed that the total output current remains stable and constant over time at **0.2417 A,** a significantly higher value compared to a single module. This increase is attributed to the parallel connection which aggregates the current from each individual string into a common busbar. Mathematically, this is expressed as $I_{total} = \sum_{i=1}^{NP} Ii.$ [23]. This configuration allows for an increase in the total system power by raising the overall current while maintaining a low voltage level compatible with inverter specifications. Furthermore, this behavior ensures power flow continuity even if one branch fails, thereby enhancing the operational reliability of the system [34].

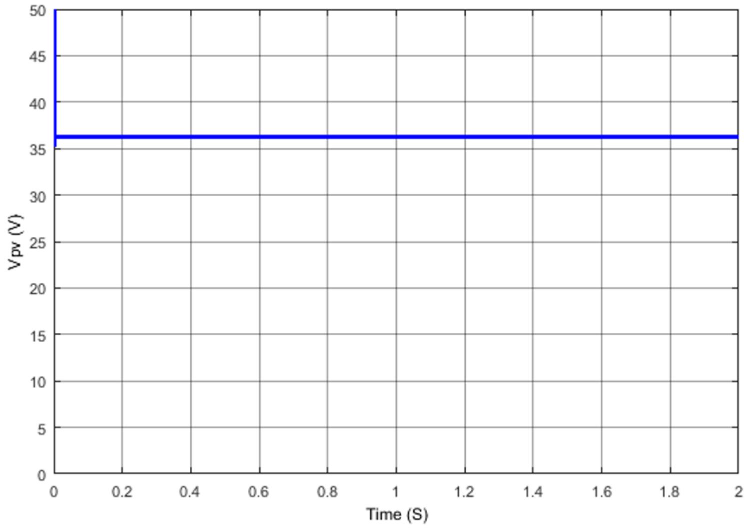


Figure II.10: PV array output current response under parallel configuration (3 single modules).

Upon examining the voltage curve (V) from the parallel connection simulation, it is observed that the value remains identical to that of a single module ($V_{\text{array}} = V_{\text{module}}$), appearing as a stable horizontal line over time. In a parallel circuit, all modules are connected across two common nodes, which forces a uniform potential difference across all branches. This physical behavoir prevents voltage accumulation (unlike series connection) supplied to the inverter. [27], notes that the voltage stability in this configuration is a technical advantage when the goal is to increase the array's current without necessitating high-voltage protection specifications of the plant.

I.2.2.4 Series-Parallel Configuration:

The series-pqrqllel configuration represents the practical of both connection types to achieve the required desing power levels. In this setup, modules are grouped into ″strings″ connected in series to boost voltage, and these strings are then interconnected in parallel to increase the total current. [28], this approach is the most prevalentin PV plants as it provides the necessary flexibility to match the specifications of central inverters.

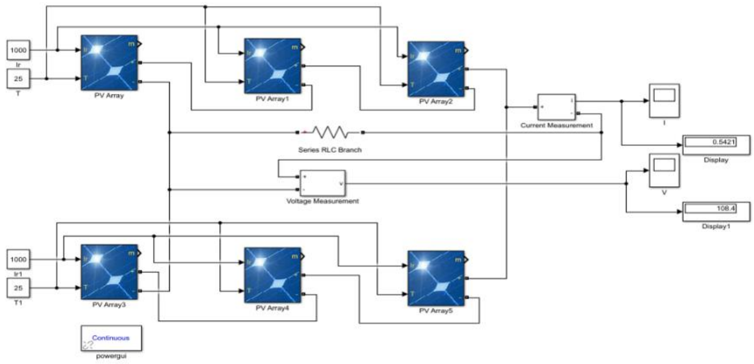


Figure II.11**:** The diagram illustrated the series-parallel configuration.

The series-parallel connected configuration is the most widely adopted setup in large-scale PV plants. It integrates series-connected groups to elevate voltage with parallel-connected groups to increase current [23]. This design facilitates reaching the target power output while ensuring system compatibility with the inverter's input voltage range. The schematic illustrates an array consisting of several strings containing several modules linked in series. These strings are then interconnected at a common busbar. This structure simultaneously combines voltage summation and current aggregation.

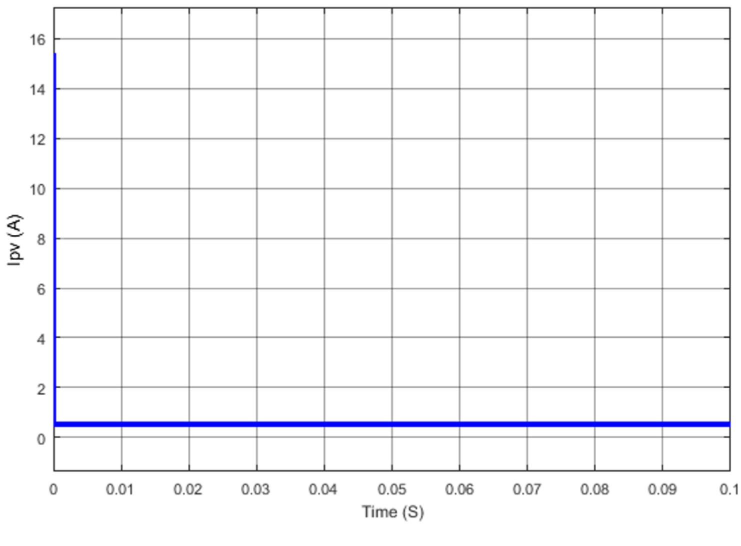


Figure II.12: PV array output voltage response under hybrid series-parallel configuration (3*2 modules).

In this hybrid architectural configuration the total operational voltage is amplified through the multiplication of series-connected modules (N_s), while the baseline capacity of the array output current scales proportionally with the number of parallel-connected strings (N_p). This mathematical integration ensures that steady-state outputs effectively stabilise at optimal operational levels, guaranteeing balanced efficiency and adequate system power. According to the foundational study by [23], this specific structural arrangement represents the most optimal configuration for large-scale array designs, as it successfully elevates the operational DC voltage which in turn reduces localised ohmic distribution power losses ($P_{loss}=I^2R$), within the system transmission cabling. Furthermore, this arrangement delivers reliable continuous current flow even under heavier loading conditions, while offering remarkable superior system resilience against partial shading abnormalities when evaluuated against standard simple series topologies [34].

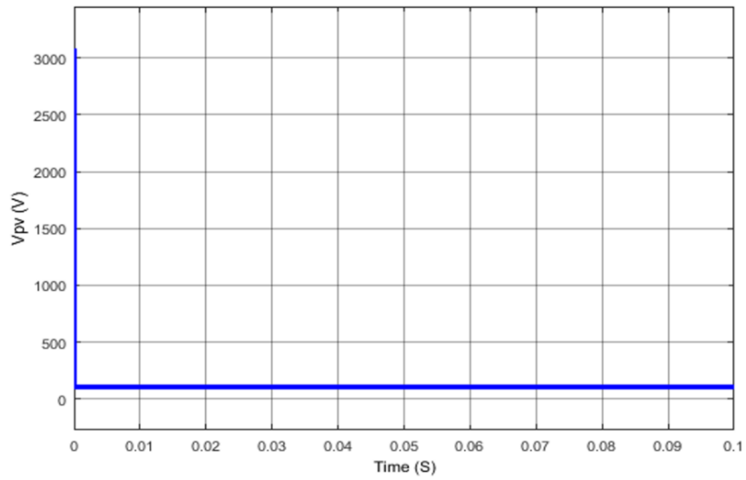


Figure II.13**:** PV array output voltage response under series-parallel configuration (3*2 modules).

This stabilised voltage profile successfully demonstrates the ideal voltage matching achieved across the series-connected modules within the independent strings of the array. Concurrently, the parallel interconnection of these strings efficiently scales the cumulative current capacity. As emphasised by modern power electronics literature, establishing this specific electrical equilibrium ensures systematic balance throughout the system's networks, verifying that the number of modules grouped in series (N_s) properly escalates the DC operating terminal voltage to satisfy input inverter requirements [23]. This structural arrangement minimises transmission thermal losses ($P_{loss} = I^2R$) while maintaining a robust barrier against sudden localised power fluctuations and microclimatic environmental dynamics. To provide a comprehensive assessment, this section establishes a quantitative evaluation of the photovoltaic network, with environmental constraints strictly fixed at Standard Test Conditions (STC: Irradiance G = 1000W/m^2, Temperature T = 25°C) to guarantee an accurate, unbiased comparative benchmark. Table (II.3) demonstrates the numerical transitions and electrical evolution recorded when moving from a simple configuration into the integrated hybrid series-parallel topology.

Table II.3: Extracted simulation electrical parameters under fixed STC conditions

Configuration	Irradiance(G) / Temp(C)	Total Voltage(V)	Total Current(I)	Total Power(P)
Series Connection	1000W/m^2 / 25°C	107.6 V	0.7174A	≈ 77.22 KW
Parallel Connection	1000W/m^2/ 25°C	36.25 V	0.2417 A	≈ 08.76 KW
Series-Parallel(Hybrid)	1000W/m^2/ 25°C	108.4V	0.5421A	≈ 58.76 KW

II.2.3 Environmental Input:

The surrounding atmosphere intrinsically links the operational efficiency of any photovoltaic architecture, necessitating a precise model of environmental inputs. In this simulation, irradiance is treated as a time-varying stimulus rather than a static equilibrium; this dynamic representation enables the observation of the system's adaptability during sudden solar transitions, a methodology emphasized by duffie and beckman [25]. Parallelly, the thermal behavior of the array is integrated through standard correction factors, acknowledging that temperature fluctuations act as a dominant constraint on the cell's voltage ceiling. To reflect the inherent unpredictability of real-world hardware, the model further introduces stochastic noise and quantisation error layers [26]**,** embedding these disturbances crucial for validating the controller's robustness, as it forces the system to maintain stability despite the persistent presence of sensor-born irregularities and environmental turbulence.

II.3 Simulation of PV Faults:

The PV array architecture utilized in this study involves a 2*3 series-parallel topology. This setup was subjected to non-uniform solar exposure to analyze the impact of partial shading on power output. Specifically, different irradiance inputs were applied to the six sub-modules to replicate complex shading patterns. The inclusion of bypass diodes in the model is critical, as they provide an alternative path for the current, leading to the emergence of multiple power peaks in the characteristic curves. The simulation employs a controlled voltage surce and a ramp signal to sweep the entire voltage range, ensuring the capture of all local and global maxima.

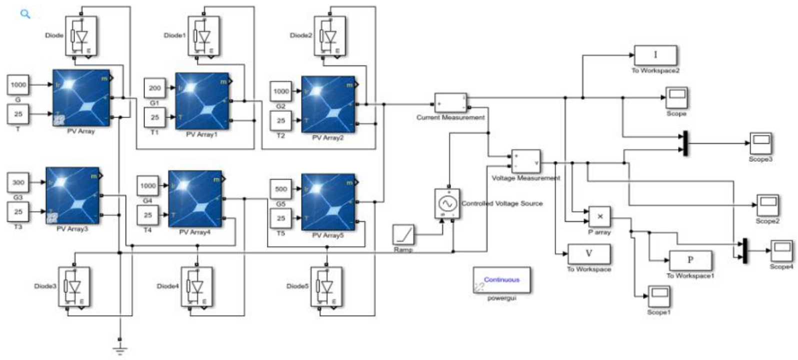


Figure II.14**:** Simulink model of PV array configuration under shading conditions with bypass diodes.

II.3.1 Partial Shading Simulation:

2.3.1.1 Effect on I-V Curve:

Once the simulation framework was established and the partial shading scenarios were defined, the performance characteristics of the PV array were evaluated. According to [23]. Modelling PV arrays under non-uniform conditions is essential to understand the non-linear behavior of the system. This section analyses the deviations in the I-V and P-V profiles, where the integration of bypass diodes is crucial for protecting modules from hot-spot effects, which lead to the formation of staircase patterns and multiple power peaks [29] The following figures illustrate these electrical phenomena, providing a basis for comparing the array's output under various mismatched irradiance levels [31].

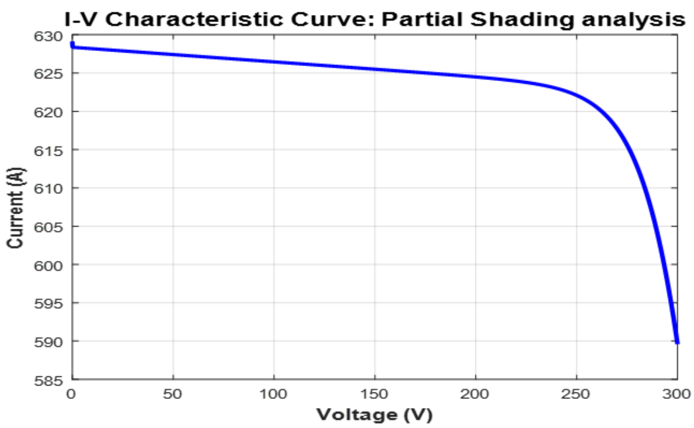


Figure II.15: I-V characteristic curve partial-shading analysis.

The obtained I-V characteristic (Figure II.15) reveals a significant deviation from the standard single-peak curves under uniform irradiance. Due to the mismatched irradiance levels across the six modules, the curve exhibits a staircase profile.[29] this phenomenon occurs because the bypass diodes become forward-biased to provide an alternative path for the current, preventing the shaded modules from acting as loads. Consequently, the total current drops at specific voltage thresholds, creating multiple steps that correspond to the different irradiance zones defined in the simulation (200, 300, 500, and 1000W/m^2). This behavior supports the findings presented in reference[23]**.** Regarding the non-linear response of PV strings under partial shading conditions.

II.3.1.2 Effect on P-V Curves:

The analysis of power-voltage (P-V) characteristics under non-uniform irradiance conditions is essential for understanding PV array performance. In this study, an array of six modules configured in two parallel strings was evaluated. The application of partial shading across different module levels complicates the electrical characteristics, resulting in the emergence of multiple Maximum Power Points (MPPs) [23]. This physical behavior presents significant challenges for conventional MPPT algorithms, necessitating a precise analysis of these peak locations to ensure maximum energy extraction and minimise losses associated with bypass diode activation [24].

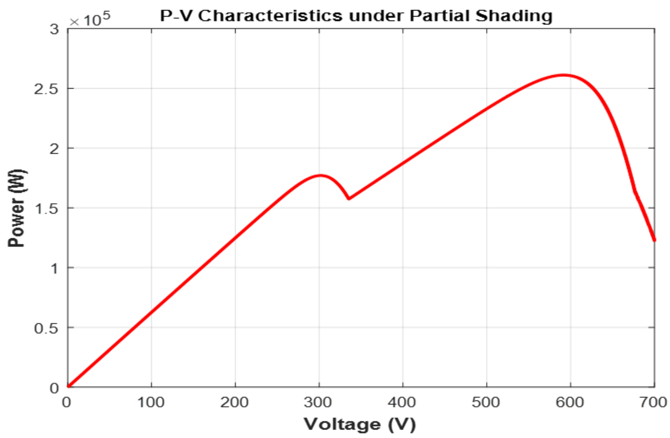


Figure II.16: P-V Characteristic Curves Partial Shading Analysis.

The simulation results for the P-V characteristic curve clearly demonstrate the impact of partial shading, characterized by the emergence of multiple Maximum Power Points (*MPPs*). According to Patel and Agarwal [23], this physical phenomenon occurs due to the activation of bypass diodes to protect shaded cells from thermal stress, which causes thecurve to split into multiple peaks. The results show a local Maximum (*LMPP*) near 300V and a Global Maximum (*GMPP*) near 600V. This aligns with the findings of Ishaque and Salam [24] regarding the complexity of the tracking process under such conditions, where advanced algorithms are essential to distinguish the global peak form local maxima and ensure optimal energy harvesting.

II.3.2 Series Resistance Fault:

The mathematical representation of the series resistance fault is implemented by modifying the equivalent circuit of the PV cell. As shown in Figure II.17**,** the resistance *Rs* is connected in series. With the cell's output, representing the losses in the metal contacts and interconnections [26]**.**

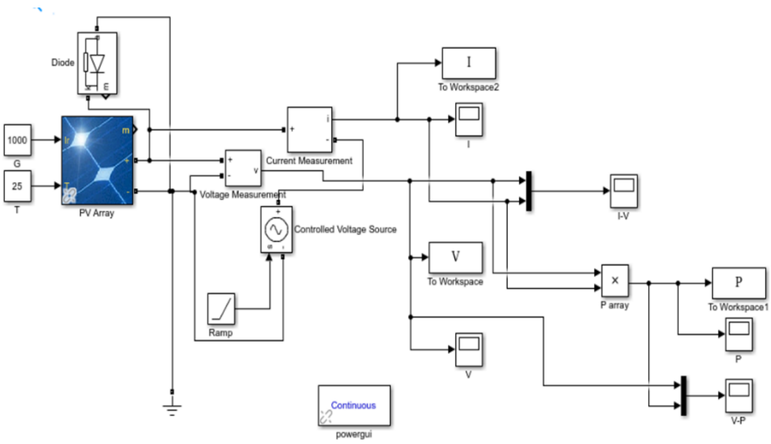


Figure II.17**:** Simulink model of the PV cell equivalent circuit, illustrating series resistance.

This section investigates the impact of series resistance (R_s) faults on the PV array performance. A rise in series resistance is a common problem caused by corrosion of the interconnectors or the ageing of the module's physical parts [25]. This degradation directly affects the slope of the I-V curve in the high-voltage region, leading to a significant reduction in the Fill Factor (FF) and the overall system power output, even under uniform irradiance conditions [26]. (Figure II.17) As shown in

II.3.2.1 I-V Characteristics: Series Resistance Fault Analysis:

A series resistance fault increases the internal resistance of the photovoltaic (PV) module, causing a significant reduction in output current and maximum power. As a result, the I–V curve becomes flatter near the open-circuit voltage region, leading to decreased efficiency and energy production.

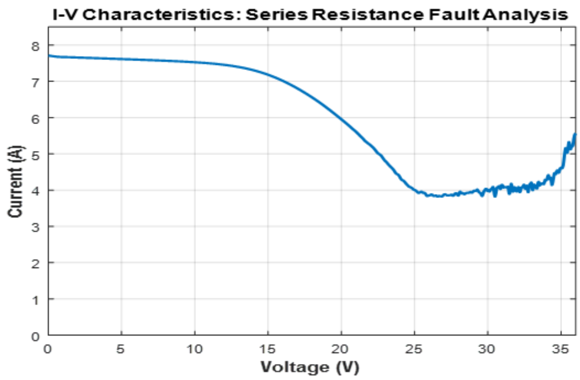


Figure II.18: I-V characteristics of the PV array under different series resistance (R_s) values.

The series resistance (Rs) is a critical internal parameter that directly influences the performance of the PV array. This resistance primarily arises from the contact resistance between the metal grid and the semiconductor, as well as the bulk resistance of the cell materials [26].

Under normal conditions, *Rs* is minimal; however, faults such as cable corrosion, solder bond failure, or thermal cycling can significantly increase its value [27].

As illustrated in Figure II.18, an increase in *Rs* leads to a notable change in the slope of the I-V curve, specifically in the region near the open-circuit voltage (V_{0c}) [24].

Unlike other faults, an increased *Rs* does not significantly alter the short-circuit current (I_{SC}), but it causes a ‘flattening’ of the curve’s knee, which drastically reduces the fill factor (FF) and the overall power conversion efficiency [35,34].

The simulation results obtained in this study (Figure II.18) are consistent with the single-diode mathematical models presented in the literature, confirming that power is dissipated as heat due to the elevated resistance [27].

II.3.2.2 P-V Characteristics of the PV Array:

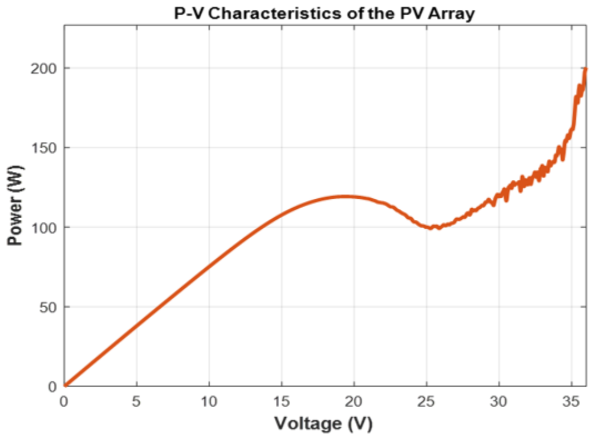


Figure II.19: P-V Characteristics of the PV Array under Series Resistance Fault.

The generated curves clearly demonstrate the impact of series resistance (Rs) on a photovoltaic array's efficiency. As illustrated in Figure II.18**,** an upward trend in *Rs* values commonly associated with contact degradation or the ageing of interconnections leads to a substantial decline in the output power[26]**.** Physically, this drop is linked to the increased internal voltage losses within the circuit, which manifests as terminal dissipation instead of useful electrical energy [23]**.** From a diagnostic perspective, the *P-V* characteristics revel that the 'fill factor' of the system is compromised as the resistances increases. The curve's (MPP) not only decreases in magnitude but also exhibits a change in its geomtric slop [34]. These observations confirm that monitoring in the *P-V* and *I-V* trajectories can serve as a reliable indicator for the early detection of ohmic-related faults in solar PV installations [35].

II.3.3 Shunt Resistance Fault (R_{sh}): In addition to series resistance, the performance of a photovoltaic (PV) array is significantly influnced by the shunt resistance (R_{sh}), which represents the parallel high-resistence path across the P-N junction. Ideally, R_{sh} should be infinite to prevent any internal diversion; however, in practical scenarios, a rerduction in the R_{sh} value orten occurs due to manufacturing defects or surface contamination along the cell edges [25]. This phenomenon leqds to what is known as 'leakage current modeling', zhere a postion of the photo-generated current is shunted away from the external load. The impact of a low R_{sh} value is most pronounced in the low-voltage region of the *I-V* characteristic curve, causing a noticeable in the slope of the curve near the short-circuit current (I_{sc}) point [27]. Understanding this fault is crucial for diagnosting power losses that occur even standard irradiance levels. The impactof these variations is clearly illustrated in figure(II.20).

II.3.3.1 I-V Characteristics: Healthy vs shunt faults:

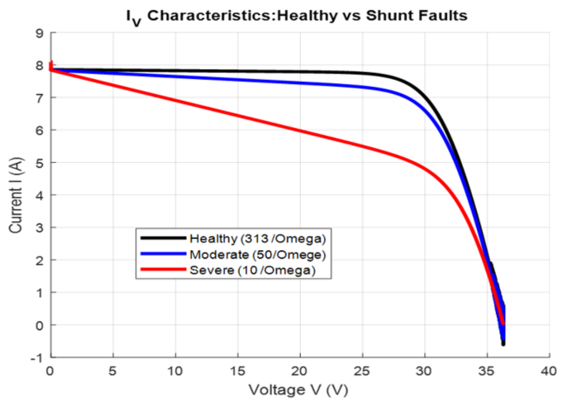


Figure II.20: Comparison *of I-V* characteristics of the PV array under various shunt resistance (R_{sh}) conditions.

The characteristic curve obtained from the simulation emphasises the degradation mechanism associated with shunt resistance variations (Rsh) [23]**.** A reduction in R_{sh} signifies the presence of alternative current paths within the solar cells, often caused by manufacturing defects or edge leakage. The results clearly demonstrate that as R_{sh} decreases from its nominal value to **10Ω**, the slope of the *I-V* curve in the voltage-source region increases significantly. This phenomenon leads to a substantial drop in the fill factor (FF) and maximum power output (P_{max}), even though the open-circuit voltage (Voc) remains relatively stable. Such behavior is a primary indicator of internal shunting faults in photovoltaic modules.

II.3.3.2 P-V Characteristics of shunt resistance (Rsh) faults:

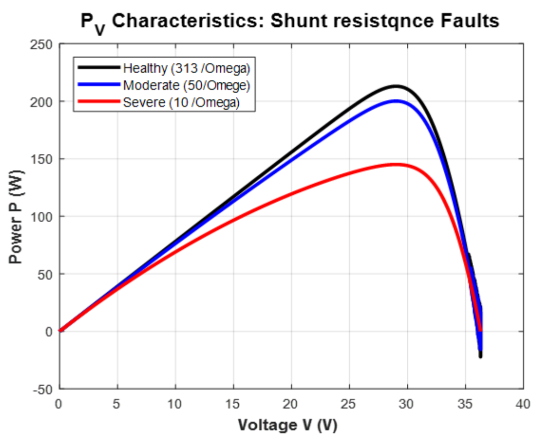


Figure II.21: Power-Voltage (P-V) characteristics show the impact of shunt resistance (R_{sh}) on the maximum power point.

The simulation results demonstrate that a significant reduction in R_{sh} leads to a substantial drop in the maximum power point (mpp) [23]. In the severe fault case (10Ω), the peak power decreased by approximately 30% compared to the healthy state, highlighting the critical impact of internal leakage currents on the overall system efficiency [24]. This behavior is consistent with the theoretical models of PV cell degradation, where low shunt resistance represents manufacturing defects or potential-induced degradation (PID) [25].

II.3.4 Bypass Diode Malfunction:

Bypass diodes are integrated into photovoltaic (PV) modules as a primary protection mechanism to mitigate the localised overheating caused by partial shading [23]. These components function by providing a low-impedance shunt path, allowing the string current to bypass shaded cells that would otherwise exhibit high resistance [24]. The simulation model developed in this study, as shown in the figure above, specifically investigates the impact of bypass diode failure, such as open-circuit malfunction often induced by thermal overstress [25]. Under these fault conditions, the current is forced to flow through the shaded cells, transforming them into dissipative loads that generate excessive heat. This process leads to the formation of 'hot spots', which significantly compromise the module's structural and functional integrity [26].

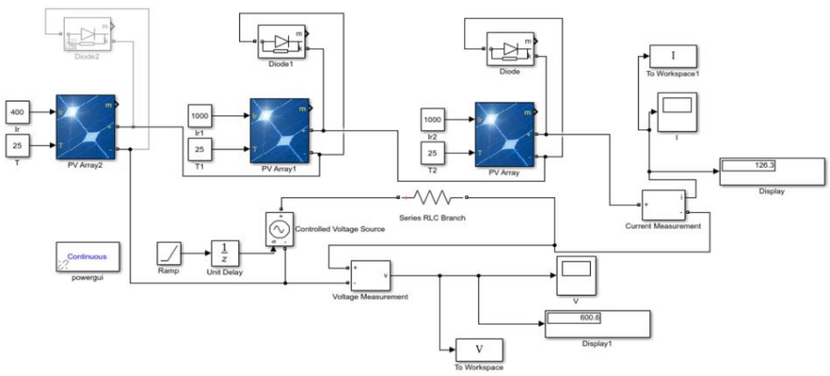


Figure II.22: Simulink model of the 3-panel PV array configuration under bypass diode failure.

II.3.4.1 Simulation Results and Performance Analysis:

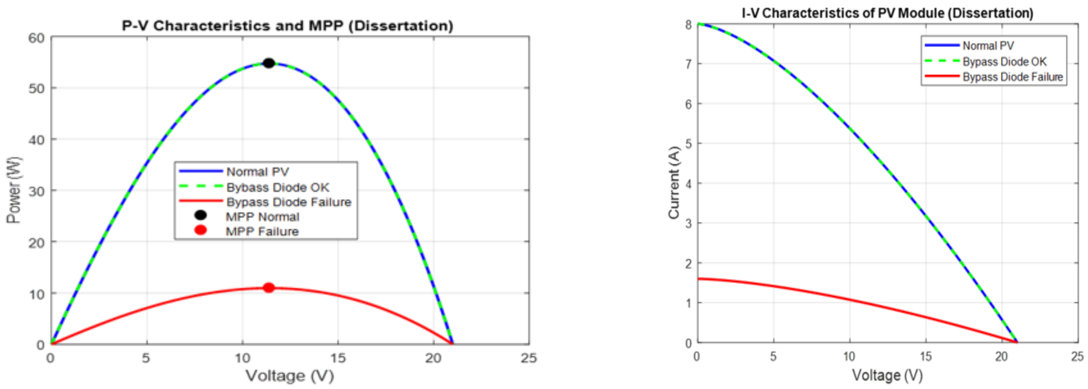


Figure II.23: Combined I-V and P-V characteristics showcasing the impact of bypass diode failure.

II.3.4.2 Quantitative Analysis of Simulation Results:

The comparative analysis between healthy and faulty states reveals the following:

- **I-V Characteristic Analysis:** The output current in the faulty states exhibits a severe reduction, dropping from a nominal **8 A** to approximately **2** A [24]. This occurs because the shaded cell, lacking an operational bypass path, acts as a high-resistance bottleneck for the entire series string [25].
- **P-V Characteristic and MPP Degradation:** The power-voltage curves illustrate a catastrophic decline in the Maximum Power Point (MPP), which falls from nearly **55W** in a healthy configuration to approximately **13W** during a diode failure [27].
- **Efficiency and Reliability Implications:** Such a failure leads to a total power loss exceeding 70%, demonstrating that bypass diode malfunctions are only a major source of energy yield loss but also a significant reliability risk [23, 25]. The resulting power dissipation within the shadedcells confirms the high probability of ″Hot Spot″ formation, which leads to permanent module damage [26].

II.3.5 Open-Circuit and Short- Circuit Faults:

Beyond shading and resistive degradation, PV array are susceptible to severe electrical failures, primarily categorized as open-circuit faults [23]. An open-circuit fault occurs due to the unintended disconnection of a module or string, often caused dy mechanical failure in interconnects or cable damage [24]. This leads to acomplete interruption of current flow through affected branch, resulting in a total loss of power from that segment [25]. Conversely, a short-circuit fault arises when an abnormal low-resistance path is established, effectively bridging the module terminals or bypass diode [26]. Such faults car be triggered by insulation failure, lightning strikes, or severe moisture ingress [23]. These conditions are critical not only because of the immediate power loss but also due to the significant electrical and safety consequences, including the risk of electric arcs and permanent component destruction [23, 27]. In this section, ze investigate the simulation of these faults to analyze their impacton array's electrical output.

II.3.5.1 Reference Case (Healthy State Simulation) :

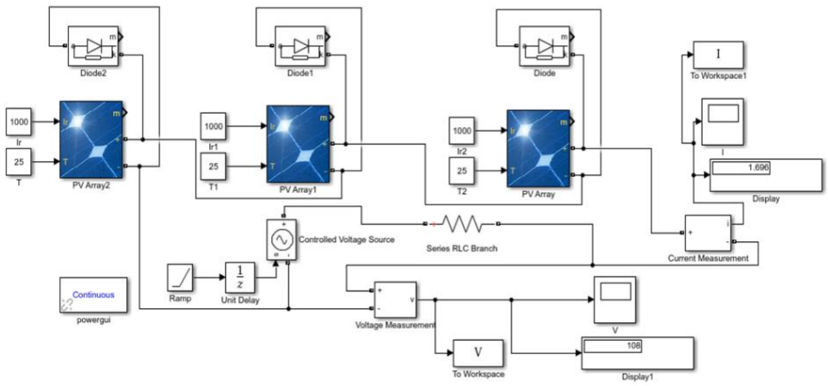


Figure II.24: Reference Case (Healthy State Simulation).

II.3.5.2 I-V and P-V Characteristics of the PV array in the Healthy state (Reference Case):

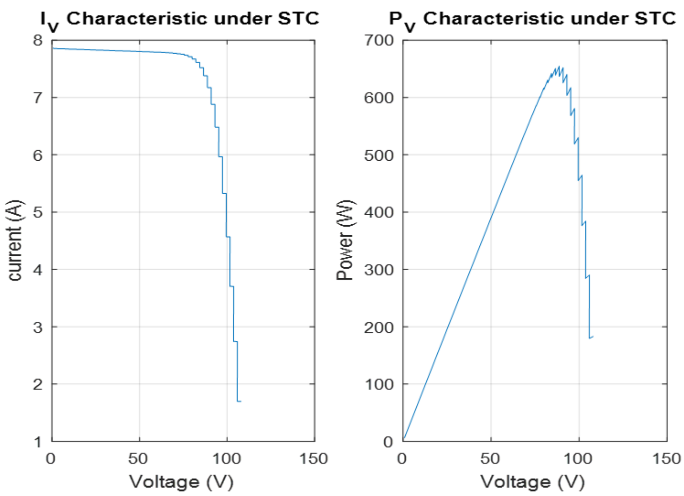


Figure II.25: I-V and P-V Characteristics of the PV array in the Healthy state (Reference Case)

The simulink results in Figure (II.25) show that the array produces a short-circuit current (*Isc*) of **7.85A** and on open-circuit voltage (*Voc*) of approximately **108.9V.** These values perfectly match the theoretical specifications of the three series-connected modules, confirming the accuracy of the established model before introducing any faults.

II.3.5.3 Open-circuit Fault Simulation :

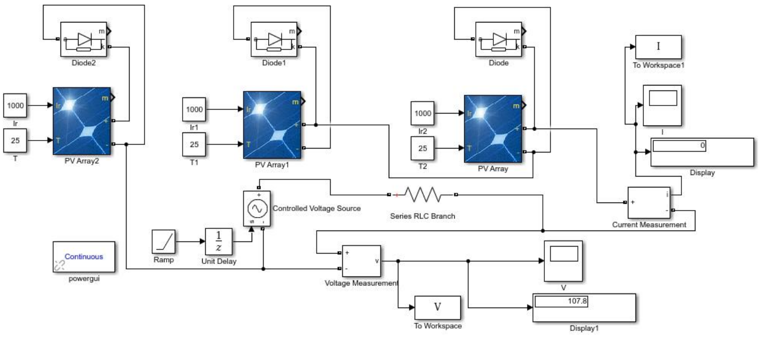


Figure II.26: Open-circuit Fault Simulation.

II.3.5.4 I-V and P-V Characteristics of the PV array in Open-circuit fault:

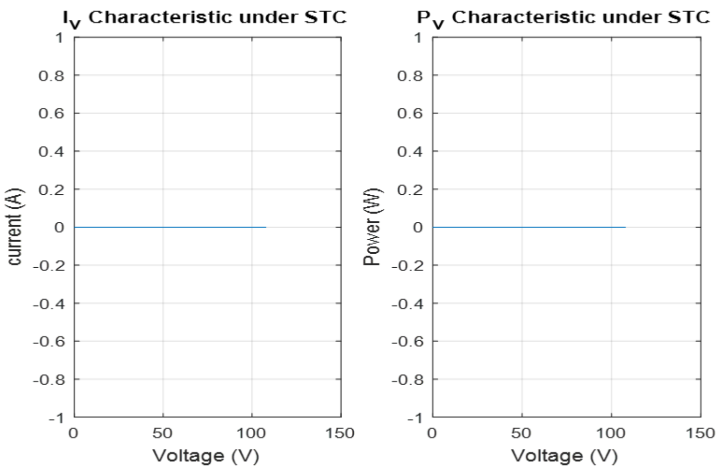


Figure II.27: I-V and P-V Characteristics of the PV array in Open-circuit fault.

The results obtained in Figure II.27 clearly demonstrate the impact of an open-circuit fault on the PV array's performance. As the connection between the modules is interrupted, the current flow drops to **zero,** resuling in a completeloss of power generation across the entire string. This confirms that even a single point or failure in a series-connected array can disable the whole system.

Note: By comparing the results of the Healthy state (figure II.25)with the Open-Circuit Fault (Figure II.27), a total power loss of 100% is observed. While the system was operating at its maximum efficiency of approximately 640W under normal conditions, the interruption of the electrical path led to a complete drop in current (I=0), disabling the energy harvesting process entirely. This highhlight the cirtical impact of interconnect failuresin series-connected PV arrays.

II.3.5.5 Simulation Results and discussion of Short-Cricuit Fault:

To evaluate the operational behavior of the 3*2 PV array under severe fault environments, a short-cricuit fault was introduced into of the parallel-connected branches. The resulting dynamic curve of the array's total terminal current and output voltage are graphically represented in figure (II.28).

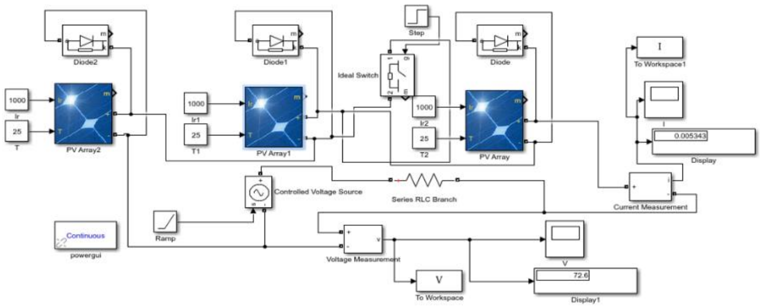


Figure II.28: PV array current and voltage trasient responses under a localized short-cricuit.

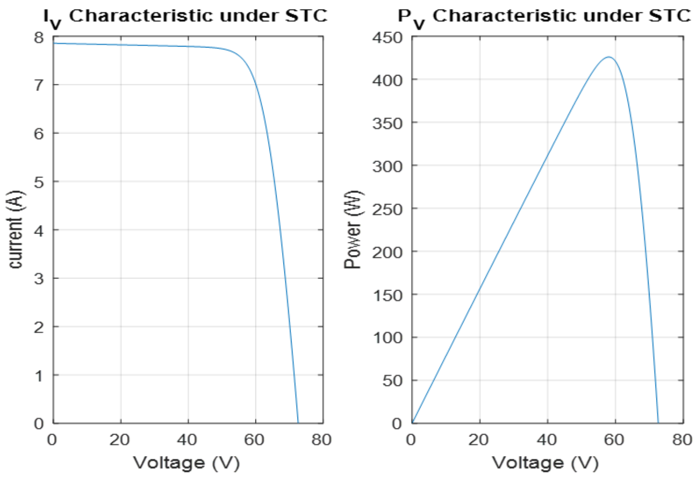


Figure II.29**:** I-V and P-V Characteristics of the PV array in Short-Circuit fault.

The numerical results extracted from the simulationreveal critical deviations from the healthy reference. Upon the inception of the short-circuit condition, the operating terminal voltage of the affected string instanly collapses toward ziro, dropping the overall configuration's aggregated output voltage. Conversely, due to the elimination of the internal load resistance wiyhin the fault path [32]. A short-circuit fault within a mult-string PV infrasttucture induces a severe current mismatch between the parallel branches, forcing a non-uniform operating point across the entire grid. While the remaining healthy string attmppts its nominal current generation, the shorted path experiences high stress. This unbalance distorts the combined $I-V$ characteristic profile and causes the overall power output to degrade substantially. Futhermore, as verified by [27], the massive voltage drop coupled with localized current aggregation under short-cricuit conditions can cause the operating to drift completely away from the Maximum point Tracker (MPPT) limits. This conrirms that unaddressed short-circuit conditions not only diminish immediate energetic ields but alsoth threaten system safety by promoting structural hot-spots across the surrounding cell encapsulations.

II.4 Data acquisition from Simulation:

The development of an intelligent fault diagnostic system based on Artificial intellgence (A) algorithems necessitates the establishment of a robust, comprehensive, and well-structured database. In this study, the required data is systematically harvested from the developed MATLAB/Simulink environment across various operating regimes, including normal conditions and pre-defined fault scenarios (such as short-cicuit, open-circuit, and partial shading). This section elucidates the sequential methodology employed to extract key electrical signatures, compute specialized performance factors, derive analytical features, and structure the final dataset for the subsequent Al classification phase.

II.4.1 Extraction of key Electrical Features:

The foundational step in the data acquisition pipeline involes monitoring and recording the fundamental electrical parameters directly from the simulation workspace during transient and steady-state operations. For each simulation instance, three synchronized primary vectors are extracted as functions of time:

1. **Array Terminal Voltage (V_{pu}):** Represents the aggregated potential difference across the configuration's common busbars.
2. **Array Output Current (I_{pv}):** Reflects the total current flowing from the combined parallel strings into the external load.
3. **Generated Power (P_{pv}):** computed continuously within the model as the direct product of the instantanrous voltage and current, expressed as:

$Ppv(t) = Vpv(t) * Ipv(t)$, utilizing raw time-domain electrical vectors directly from the simulation outputs helps capture the macro-level deviations induced by electrical faults. These foundational variables serve as the core indicators from which more complex diagnostic features are subsequently derived [34].

II.4.2 Fill Factor (FF) Calculation:

To evaluate the degree of electrical degradation and profile distortion caused by different fault categories, the Fill Factor (*FF*) is continuously evaluated. The *FF* is a dimensionless metric that quantifies the overall quality and performance efficiency of the PV plant by comparing the maximum attainable power against the theoretical upper limits defined by the Open-Circuit Voltage (V_{oc}) and Short-Circuit Current (I_{sc}). Mathematically, it is defined as:

$$FF = \frac{Pmp}{Voc * Isc} = \frac{Vmp * Imp}{Voc * Isc} \tag{5.II}$$

Where V_{mp} and I_{mp} represent the voltage and current at the Maximum Power Point (MPP), sespectivety[30]. Localized anomalies such, as partial shading or intro-string short-circuits_introduce severe mismatches that heavily distort the *I–V* characteristic curve, leading to a noticeable drop in the Fill Factor compared to the healthy reference baseline. Consequently, including the *FF* as a feature drastically enhances the classification accuracy of the Al model.

II.4.3 Derivative of the I – V Curve *(dI / dV)*:

While static electrical parameters provide substantial diagnostic insight, they are often insufficient for distinguisging between complex overlapping anomalies, such as partial shding and multi-module resistance faults. To circumvent this limitation, the first mathematical derivative of the current with respect to voltage *(dI / dV)* is calculated across the operational trajectory. The integration of the *dI / dV* profile allows the diagnostic fromwork to mathematically identify the slope transitions and the presence of multiple local peaks (inflection points) caused by the activation of bypass diode under non-uniform irradiance profiles **(Chouder & Silvestre, 2010).** This derivative serves as a highly sensitive geometric feature that isolates electrical faults from environmental fluctuations.

II.4.4 Extraction of Statistical Features (Mean, RMS, Energy):

Because processing raw, high-resolution time-series data maximizes computational overhead and can lead to overfitting in machine learning algorithms, data reduction is achieved by transforming the continuous electrical signals into discrete statistical attributes. For every simulated event, three distinct statistial markers are computed from the voltage, current, and power profiles:

1. **Mean Value (μ):** Provides the central tendency of the electricaloutput over the simulation window (*T*), defined as

$$\mu = \frac{1}{T}\int_0^T f(t)dt \qquad \text{(II.6)}$$

2. **Root Mean square (RMS):** Quantifies the effective magnitud and captures the severe fluctuations or ripple effects induced by transient fault:

$$RMS = \sqrt{\frac{1}{T}\int_0^T [f(t)]^2 dt} \qquad \text{(II.7)}$$

3. **Signal Energy (E):** Measures the integral area under the power curve to reflect the cumulative energetic yield and quantify the total energy loss caused by the operational anomaly:

$$E = \int_0^T Ppu(t)dt \qquad \text{(II.8)}$$

The deployement of these statistical descriptors ensures that the input space is compact yet rich in diagnostic information, which standerd pre-processing protocols established in modern solar fault classification literature [27].

II.4.5 Dataset Structure for Al Classification:

The final phase of the data acquisition task involves compiling the extracted features into a unified matrix structure compatible with machine leraining and Al classification toolboxes. Organized into an M*N matrix, where each row (M) represents an independent simulation trial (a specific operational samble), and each column (N) corresponds to a specific extracted attribute (Feature). The input vector for a single data instance is structured as follaws:

$$X = [Mean(V), RMS(V), Mean(I), RMS(I), FF, dI/dV, Enrgy]. \qquad \text{(II.9)}$$

To facilitate supervised learning, a deeicated ″Label″ or ″Target″ column (Y) IS appended to the end of each row. This label utilizes numerical indexing to define the true operational state of the array, for instance:

- **Y** = 0: Healthy State (Normal Operation).
- **Y** = 1: Short-Circuit Fault.
- **Y** = 2: Open-Circuit Fault.

- **Y** = 3: Partial Shading Condition.

This structured dataset provides the formal fromework necessary to train, validate, and test the machine learning classifiers responsible for automated fault detection.

II.5 Data Per-processing for Machine Learning:

Once the electrical signatures and statistical descriptors are harvested from the simulation environment, they must undergo a rigorous pre-processing phase before being introduced the artificial intellgence (AI) classifier. Raw simulation data frequently contains structural disparities, scales of magnitude variations, and transient noise that can severely suboptimize machine learning convergence. This section details the systematic pre-processing pipeline executed to refine, normalize, filter, and select the most deterministic features, concluding with the generation of the integrated benchmark dataset.

II.5.1 Normalization and Scaling:

In the collcted dataset, the extracted features exhibit vast numerical dispatites; for instance, the array terminal voltage operates of tens or hundreds of volts, whereas the output current operates in small fractions of an ampere. If left unscaled, features with larger magnitudes will dominate the objective function of the AI classifier, rendering smaller but equally critical parameters redundant. To mitigate this geometric bias, a Min-Max normalization technique is deployed to scale all feature bounds into a uniform continuous closed interval of [23, 24]. The mathematical transformation is governed by the following exressio

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad \text{(II.10)}$$

Where: X is the original feature value, Xmin and Xmax represent the minimum and maximum boundaries af that specific feature across the entir dataset, and X$norm$ represents the resulting standardized value [30]. Executing unform scaling accelerates the gradient descent convergence during the training phase and elevates the classification boundary resolution.

II.5.2 Noise Filtering Techniques:

During severe operational transitions, such as the inception of a short-circuit fault or rapid multi-module irradiance changes under partial shading, high-frequency transients and numerical fluctuations naturally occur within the continuous simulation vectors. To prevent these localized anomalies from degrading the training path, digital smoothing filters, specifically moving average algorithms, are applied to the time-series profiles before feature extraction [32]. Removing non-characteristic transient noise ensures that the classifier focuses entirely on steady-state electrical profiles, thereby eliminating flse-positive classifications during dynamic system operations.

II.5.3 Handling Missing and Outlier Data:

Numerical stiffness within the ordinary differential equation (ODE) solvers of MATLAB/Simulink can occasionally introduce mathematical outliers or unrecorded data gaps (NaN values) during severe fault conditions. To preserve the statistical equilibrium of the matrix, an automated screening protocol is developed. Outliers lying beyond three standard deviations from the feature mean are detected and substituted using linear interpolation based on adjacent homogenous samples, ensuring data continuity and preventing alogrithmic compilation fialures during supervised training [27].

II.5.4 Feature Selection Strategies :

To reduce dimensionality and eliminate redundant information that could lead to overfitting, a systematic feature selection strategy is conducted. High dimensionality increases computational overhead and can cloud the classification boundaries. By assessing the correlation matrix and utilizing evaluation metrics, only the features dispiaying the highest variance and lowest mutual dependency are preserved. This technical screening balances high training speed with maximum diagnostic resolution, ensuring a lean yet highly informative input vector space [27].

II.6 Benchmark Dataset Integration:

To optimize the diagnostic capability, scalability, and generalization barries of the proapective artificial intelligence (AI) fromework, the fault platform must not soley rely on synthetically generated data. Methodological reliance on single-source simulation environments often restricts model performance when exposed to industrial, non-linear solar behaviors. To overcome this limitation, this section details the systamic integration of the simulated electrical parameters with verified, publicly available open-source empirical fromeworks, establishing a hybrid benchmark database designed for robust classifier training.

II.6.1 Characterization of the YOCASOL Benchmark Dataset:

A core component of the experimental validation process involves the incorporation of standard benchmark open-source repositories, primarily the YOCASOL data platform. The YOCASOL environment delivers a comprehensive and highly curated dataset composed of measured physical parameters harvested from practical, grid-connected solar photovoltaicconpremises configurations under various climatic dynamics and degradation faults [36], utilizing standard field-measured datasets like like YOCASOL provides the machine learning classifier with highly relevant operating parameters, including localized cell temperatures, ambient insulation fluctuations, and true degradation curves. Introducing this standardized baseline allows for impartial verification and provides a definitive cross-comparison platform to assess the accuracy of modern automated diagnostic.

II.6.2 Synthesis of Simulation Matrices and Field-Measured Data:

To capture both the numerical stability af mathematical modelsand the stochastic behavior of physical fields, a systematic data fusion matrix is executed. The high-resolution feature vectors derived from the MATLAB/Simulink must-module configurations (3*2 matrix) are structural, synchronized, and blended with the empirical open-source vectors from the YOCASOL dataset. This combination maps the idealized theoretical fault behaviors alongside true oprational features, effectively filtering out model-induced baseline biases. As established in recent solar data-driven architectures, merging mult-domain data streams enriches the decision-making boundaries of supervised learining classifiers, smoothing the migration from software simulation to real-world industrial embedded microcontrolle[30].

II.6.3 Diversity Strategies for Robust Machine Learning Training :

A common pitfall in deploying fault classification models is algorithmic overfitting, wherein a neural network or classifier scores high validation metrics on simulation parameters mus fails when exposed to field environments. To guarantee high immunity against this operational drift, a strict data diversity protocol is enforced during the matrix construction phase. By compiling extensive samples that systematically span helerogeneous irradiance scales, divergent ambient thermal states, and overlapping muti-fault parameters, the cross-entropy profile of the data is maximized. This strategic diversification yields an exceptionally robust training matrix, enabling the classifier to extract deep, resilient fault signatures while ignoring environmental disturbances and localized sensory noises[27].

II.7 Validation of Simulation Models:

Before deploying the acquired datasets into the artifical intelligence classification environment, the mathematical and physical fidelity of the developed software prototype must be rigorously verified. Verification processes ensure that the current-voltage (I-V) and power-voltage (P-V) characteristic profiles generated in MATLAB/Simulink adequately emulate industrial engineering baselines under matching environmental stresses. This section outlines the validation protocols through direct evaluations, and final parametric adjustments implemented to optimize mathematical accuracy.

II.7.1 Comparative Analysis of Simulated and Experimental Curves:

The primary phase of verification relies on direct overlay comparison between the simulated characteristic trajectories and real-world experimental measurements recorded under identical Standard test Conditions (STC: $G = 1000$ W/m^2, $T = 25$°C). The non-linear currents and powers derived across the voltage sweep are plotted simultaneously alongside empirical reference data extracted from practical solar cells. As emphasized by [23], achieving a close visual and geometric alignment along the curve's three critical operational junctions_the short-circuit region, the maximum power point curve, and the open-circuit knee provides immediate verification of the diode circuitmodeling equations under steady-state conditions.

II.7.2 Statistical Performance Metrics (RMSE and Correlation Coefficient):

To complement the visual comparison between the simulated and experimental photovoltaic characteristics, quantitative validation was performed using two widely adopted statistical indicators: the Root Mean Square Error (RMSE) and the Pearson Correlation Coefficient (R). These metrics provide an objective assessment of the agreement between the simulated model outputs and the corresponding experimental measurements.

II.7.2.1 Root Mean Square Error (RMSE):

The Root Mean Square Error (RMSE) measures the average magnitude of the deviation between the simulated current values and the experimental observations. It quantifies the residual error by computing the square root of the mean of the squared differences between the simulated and measured currents. A lower RMSE value indicates a better fit of the simulation model to the experimental data.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_{exp,i} - I_{sim,i})^2} \tag{II.11}$$

Where:

- $I_{exp,i}$is the experimental current value at the i^{th}operating point;
- $I_{sim,i}$is the simulated current value at the i^{th}operating point;
- $N$is the total number of measurement samples.

The RMSE is expressed in the same unit as the measured variable (amperes in this case). An RMSE value approaching zero indicates excellent agreement between the simulated and experimental current profiles.

II.7.2.2 Pearson Correlation Coefficient (R):

The Pearson Correlation Coefficient (R) evaluates the degree of linear relationship between the simulated and experimental datasets. Its value ranges from −1 to +1, where a value close to +1 indicates a strong positive correlation and excellent agreement between both datasets.

$$R = \frac{\sum_{i=1}^{N}(I_{exp,i}-\bar{I}_{exp})(I_{sim,i}-\bar{I}_{sim})}{\sqrt{\sum_{i=1}^{N}(I_{exp,i}-\bar{I}_{exp})^2 \sum_{i=1}^{N}(I_{sim,i}-\bar{I}_{sim})^2}} \tag{II.13}$$

Where :

- $\bar{I}_{exp}$is the mean experimental current;
- $\bar{I}_{sim}$is the mean simulated current;
- $N$is the number of data samples.

A correlation coefficient close to 1 demonstrates that the simulated model accurately reproduces the trend and dynamic behavior of the experimental measurements. Therefore, the combined use of RMSE and the Pearson correlation coefficient provides a comprehensive evaluation of both the absolute error and the similarity of the profiles, ensuring a rigorous validation of the photovoltaic system model According [36], an

RMSE value approaching zero combined with a correlation coefficient (R) exceeding 0.99 mathematically confirms that the internal dynamics of the multi-model configuration are correctly modeled.

II.7.3 Model Adjustment and Parametric Fine-Tuning:

To minimize the calculated residual errors and eliminate slight deviations observed along the $I - V$ curve's curve inflection points, a systematic model adjustment protocol was executed. Internal solar cell parameter matrices_specifically the module series resistance (R_s), shunt parallel resistance (R_{sh}), and the diode ideality factor (a)_were iteratively fine_tuned. This parametric optimization successfully reduced the mathematical stiffness of the model, bringing the final simulation outputs into optimal alignment with real-world physical behavior, thereby finalizing a robust dataset for subsequent Al fault classification step.

II.8 Visualization of Fault Effects :

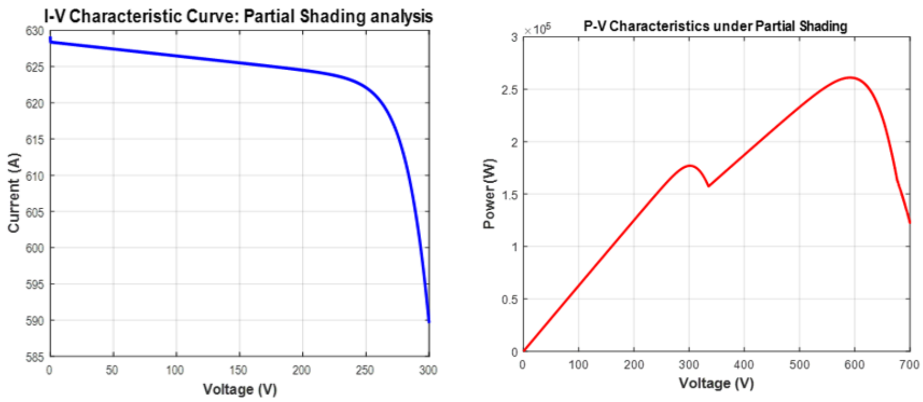


Figure II.30.a: illustrated the I-V and P-V partial shading analysis displaying multi-peak distrortions.

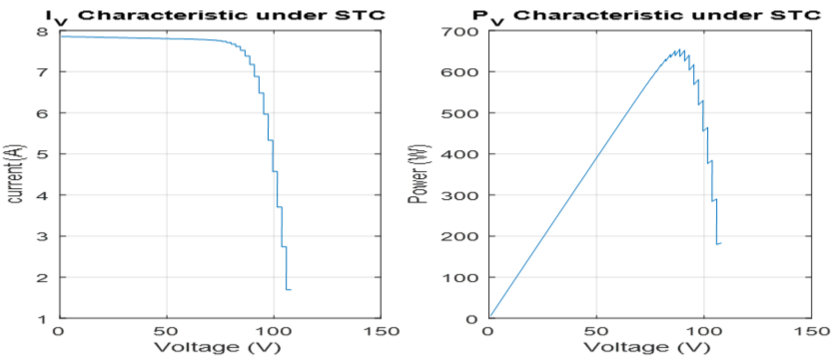


Figure II.30.b: illustrated the electrical characteristics of Healthy state reference case under Standqrd test Conditions (STC).

Analysed: The macro-elecrical behavior evaluated in Figure (30.II), establishes a critical operational baseline for distinguishing uniform environmental inputs from localized mismatch stresses. Under nominal standard operating conditions (b), the configuration tracks a predictable, single-peak current-voltage and power-voltage path, yielding a maximum power generatio limit near 640 W. As demonstrated by [1], uniform irradiance ensures that all internal cell junctions operat symmetrically without triggering bypass mechanism protections. Conversely, the introduction of localized partial shading (a) forces an immediate geometric testructuring of the characteristic profiles. Due to the activation of parallel bypass diodes protecting the shaded substrings, the curves exhibit a distinct multi-peak staircase distortion [32], this visual deformation splits the electrical response into multiple local maxima alongside a single global Maximum Power Puints (MPP), creating a complex feature space for subsequent diagnostic profiling.

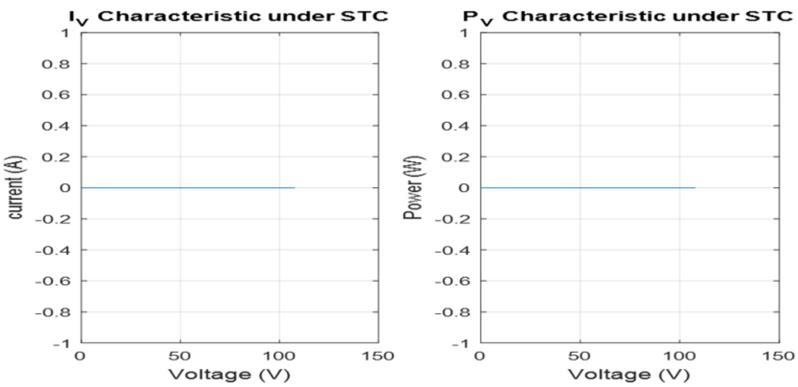


Figure (II.31.a): Dynamic electrical signatures under electrical fault stresses for the Open-Circuit fault status exhibiting total parameter collapse.

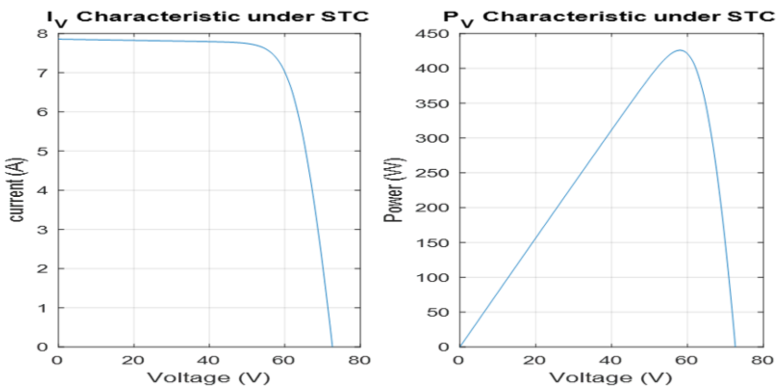


Figure (II.31.b): Dynamic electrical signatures under electrical fault stresses for the Short-Circuit fault condition under structural mismatch stress.

Analysed:

The variation in electrical parameters induced structural line faults are systematically in Figure (II.31). In the case of on open-circuit fault scenario (a), the physical continuity of the string is completely broken, preventing any electron transport across the external load terminals. Consequently,[30]. Both current and power attributes experience a total, instantaneous collapse to ziro across the entire voltage sweep range. On the alternative spctrum, the short-circuit fault topology (b) exhibits an aggressive horizontal contraction restricting the voltage coordinate axis. While the nominal short-circuit current remains temporarily sustained across un-faulted paths, the overall operational voltage capability shrinks severeiy. This distinct reduction, establishes a unique geometric 'fault signatyre' or fingerprint, capturing the exact impedance drop caused.[34]

II. 8.1 The results of faults simulation:

The results of the faults simulation illustrate in the table II. 4.

Table II.4: results of faults simulation.

Operating Condition	Peak Power (*Pmp*)	Voltage at MPP (*Vmp*)	Current at MPP (*Imp*)	Fault Severity Status
Healty Regime (STC)	≈ 854.86 W	≈ 108.9 V	≈ 7.85 A	Nominal Reverence Case
Partial Shading Fault	≈ 255 kW	≈ 590 V	Multiple Peaks	Multi-Peak Distortion
Short-Circuit Fault	≈ 425 W	≈ 52 V	≈ 8.1 A	Voltage Axis Contraction
Open-Circuit Fault	0 W	0 V	0 A	Total Parameter Collapse

II.9 Conclusion:

This chapter has systematically established a validated simulation framework essential for analyzing the macro-electrical behaviors of photovoltaic matrices under varied operational states. By executing equivalent single-diode mathematical formultion alongside a physical string arrangement implemented in MATLAB/Simulink, a reliable baseline for solar array behavior wos achieved.

The deliberate modeling of non-uniform and faulty conditions specifically localized partial shading, open-circuit paths, and module-level short-circuit carries profound diagnostic weight [30]. Accurate fault modeling is indispensable since distinct degradation modes introduce highly localized geometric deviations into the standard current-voltage (I-V) and power-voltage (P-V) trajectories. These structural anomalies capture impedance variations that remain completely hidden during nominal, clear-sky operation. Ultimately, the validation indices demonstrated that the generated dataset accurately mirror practical solar cell behaviors with high statistical fidelity [33]. This intensive data preparation phase is a mandatory prerequisite for robust Al-based classification. The filtered, normalized, and well-structured feature vectors derived from this simulation block establish the core training matrix needed to optimize and evaluate the machine learning diagnostic architectures detailed throughout Chapter III.

Chapter III:

Artificial Intelligence-Based Fault Diagnosis

III.1 Introduction:

The world is currently experiencing a surge in technological development and a growing interest in using photovoltaic panels, or PV systems, to generate electrical energy. Its uses have expanded from homes to public places, all the way to large photovoltaic stations. Under this expansion, the issue of exploring faults and diagnosing them quickly and accurately has become an inevitable necessity to maintain the efficiency of the system and the continuity of electrical energy production. Such faults include partial shading, bypass diode faults, increased series resistance, decreased shunt resistance, and others that lead to a decline in efficiency and a reduction in the system's power, potentially resulting in significant economic losses.

For this reason, there has been a trend towards using the artificial intelligence technologies that can detect faults early, analyse complex data, identify different types of faults and improve the accuracy of diagnoses in real time to improve performance and system reliability. Resorting to the use of artificial intelligence (AI) is due to the limitations of traditional methods for diagnosing and detecting faults, which often depend on manual inspection, direct electrical measurement, and thermal imaging; the difficulty of early detection; and time consumption and cost.

Among the artificial intelligence mechanisms, which in turn use several algorithms to classify faults in PV systems, the main goal is to determine the type of fault based on data extracted from the system. Like the artificial neural network (ANN), it's used to learn the relationship between the inputs and outputs of the system; the conventional neural network (CNN) is used for 2D images and data processing, like analysing thermal images of panels and hot spots; the support vector machine (SVM) is a powerful classification algorithm that relies on data separation; and the principal component analysis (PCA) is used to reduce the number of variables while retaining the most important information. There are other techniques of AI. In this chapter, we will present some AI mechanisms to diagnose some of the faults with PV panels.

III.2 Data Preparation for AI Models:

III.2.1 Dataset Formation:

When diagnosing faults in PV panels using the artificial intelligence (AI) techniques from the artificial neural networks (ANN) or convolutional neural networks (CNN) and machine learning classifiers, a well-structured and representative dataset is required to achieve accurate photovoltaic (PV) fault diagnosis. The quality of the dataset directly affects the learning capabilities, generalisation performance, and robustness of the developed diagnostic model. Therefore, careful preparation of the dataset is considered one of the most critical stages in AI-based PV fault detection systems. In this work, the dataset is constructed using both simulated PV system data generated in MATLAB/Simulink and reference datasets obtained from experimental or benchmark sources such as the YOCASOL Project. The integration of simulated and benchmark data improves the diversity of operating conditions and enhances the reliability of the trained AI model. Dataset of simulated PV system generated in MATLAB/Simulink flowing as (table III.1).

Table III.1: Illustrated dataset of simulated PV system.

Fault Type	Configuration Scale	Specific Condition	*Vpv* (V)	*Ipv* (A)	*Ppv (W)*
Normal State (Healthy)	3 PV Modules	*Rs* = 10Ω *Rsh* = 313Ω	24.0	6.20	150.0
Partial Shading	6 PV Modules	/	600	4.30	$2.6* 10^5$
Series Resistance (Rs)	1 PV Modules	*Rs* = 10Ω *Rs* = 50Ω *Rs* = 313Ω	22.0 23.0 37.0	6.20 6.10 5.70	135.0 140.0 210.0
Shunt Resistance (*Rsh*)	1 PV Modules	*Rsh* = 10Ω *Rsh* = 50Ω *Rsh* = 313Ω	24.0 23.0 36.0	6.10 5.50 3.70	145.0 125.0 135.0
Bypass Diode Failure	3 PV Modules	/	11.5	1.15	13.00

III.2.2 Labelling fault types:

Labelling fault types is the process of assigning a specific class or category to each data sample according to the fault condition it represents in table III.2.

Table III.2: illustrated the labelling of faults of PV system.

Fault types	**Label**
Normal operating	**0**
Partial shading	**1**
Rs increase	**2**
Rsh decrease	**3**
Bypass-diode failure	**4**

III.2.3 Feature Extraction:

III.2.3.1 Electrical features:

These electrical features (I, V, P, and FF) are the importance features in AI- based fault diagnosis of the PV panel, and widely used as input variables for machine learning and deep learning models such as Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Support Vector Machines (SVM), and PCA-based classifiers. Changes in these parameters provide distinctive signatures that allow AI algorithms to identify and classify different PV faults with high accuracy [37].

The output current (I) of a PV module is directly influenced by solar irradiance and reflects the ability of the module to convert solar energy into electrical energy. Current measurements are highly sensitive to faults such as partial shading, open-circuit conditions, and cell degradation. A significant reduction in current often indicates a loss of photovoltaic generation capability [37]. The PV output current can be expressed using the single-diode model as:

$$I_{pv} = I_{ph} - I_D - I_{sh} \tag{III.1}$$

The output voltage(V) is another important electrical parameter that reflects the electrical state of the PV module. Voltage variations are mainly influenced by temperature changes, load conditions, and internal electrical faults. Reduced voltage levels may indicate short-circuit faults, poor electrical connections, or increased series resistance [39]. The PV voltage is commonly represented by:

$$V_{PV} = V_{OC} - I_{PV} R_S \tag{III.2}$$

The power output (P): The electrical power generated by the PV module is calculated from the measured voltage and current values. Power is one of the most informative indicators of PV system performance because most faults directly affect energy production. Monitoring power variations allows rapid detection of abnormal operating conditions.

$$P_{pv} = V_{pv} . I_{pv} \tag{III.3}$$

The fill factor is a dimensionless parameter that evaluates the quality and efficiency of a PV module. It measures how closely the actual maximum power output approaches the theoretical maximum power defined by the open-circuit voltage and short-circuit current. The fill factor is particularly useful for identifying degradation mechanisms and internal faults that affect the shape of the current-voltage characteristic [38]. The fill factor is defined as:

$$FF = \frac{V_{mp} I_{mp}}{V_{oc} I_{sc}} \tag{III.4}$$

Fault effect on FF :

- Rs increase significantly decreases FF
- Rsh degradation reduces FF moderately.
- Partial shading creates unstable FF values.

Because fill factor combines several electrical characteristics, it is considered a powerful indicator for PV fault diagnosis.

III.2.3.2 Derived features:

The slope of the current–voltage (I–V) characteristic curve is an important derived feature for photovoltaic (PV) fault diagnosis. It provides information about the dynamic electrical behavior of the PV module and is highly sensitive to changes caused by faults, degradation, shading conditions, and variations in operating parameters [61]. The slope of the I–V curve is defined as the rate of change of current with respect to voltage:

$$S_{IV} = \frac{dI}{dV} = -\frac{I}{V} \tag{III.5}$$

For discrete measurements obtained from simulations or real-time monitoring systems, the slope can be approximated as:

$$S_{IV} \approx \frac{I_{i+1}-I_i}{V_{i+1}-V_i} \tag{III.6}$$

where $I_i$and $V_i$represent successive current and voltage measurements.

Different regions of the I–V curve is affected differently by faults, like the near short-circuit region, which is sensitive to shunt resistance faults and the leakage currents. In the near open-circuit region, the system is sensitive to increases in Rs series resistance and the ageing effects of solar cells. Therefore, changes in the slope help AI models identify curve deformations, multiple operating points, and abnormal electrical transitions [60].

III.2.3.3 Statistical features:

Statistical features are widely used in photovoltaic (PV) fault diagnosis because they provide compact and informative representations of electrical signals. These features summarize the behavior of measured parameters such as current, voltage, and power over a given observation period. Statistical indicators are particularly useful for Artificial Intelligence (AI)-based diagnostic systems because they reduce data dimensionality while preserving essential information related to the operating condition of the PV system [50].

A. **The mean:** is the value represents the average magnitude of a signal over a specific interval.

For a signal x_i, the mean is calculated as:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{III.7}$$

where: $\mu$is the mean value, $x_i$is the i^{th}sample of the signal, $N$is the total number of samples.

- Average current analysis.

- Average power monitoring.
- Stability assessment.

Faults often modify average signal values significantly.

B. The root mean square (RMS): is defined as

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{8.III}$$

C. Energy: is defined as

$$E = \sum_{i=1}^{N} x_i^2 \tag{III.9}$$

III.2.4 Feature Preprocessing:

Feature preprocessing is a crucial stage in photovoltaic (PV) fault diagnosis systems because the quality of the input data directly affects the performance of machine learning models. Raw PV data collected from sensors and monitoring systems often contain variables with different scales, missing values, noise, and redundant information. Therefore, preprocessing techniques such as normalization, missing value handling, noise filtering, and dimensionality reduction are applied before training the diagnostic model**.**

III.2.4.1 Normalizations and Scaling:

The extracted features, including current (I), voltage (V), power (P), fill factor (FF), and statistical indicators, may have significantly different numerical ranges. Machine learning algorithms are often sensitive to such variations, as features with larger magnitudes can dominate the learning process.

Normalization transforms the data into a common scale while preserving the relationships between variables. One of the most commonly used methods is Min-Max normalization:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{III.10}$$

Where: $\boldsymbol{x}$ is the original feature value, $\boldsymbol{x_{min}}$ is the minimum value of the feature, $\boldsymbol{x_{max}}$ is the maximum value of the feature, $\boldsymbol{x_{norm}}$ is the normalized value. The normalized values are typically bounded between 0 and 1.

III.2.4.2 Handling Missing Values and Noise:

PV monitoring systems may generate incomplete or noisy datasets due to sensor failures, communication errors, environmental disturbances, or data acquisition issues. Missing and corrupted data can significantly reduce the reliability of fault diagnosis models.

III.2.4.3 Missing Values Treatment:

Missing values can be handled using several approaches:

$$x_{miss} = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad \text{(III.11)}$$

- ✓ Removal of incomplete records when the amount of missing data is negligible.
- ✓ Mean imputation.
- ✓ Median imputation, which is more robust to outliers.
- ✓ Interpolation techniques for time-series PV data.

The choice of method depends on the percentage and distribution of missing observations.

III.2.4.4 Noise Reduction:

Noise filtering improves signal quality and enhances feature reliability. Common techniques include of the moving average Filter flowing as:

$$y(k) = \frac{1}{M}\sum_{i=0}^{M-1} x(k-i) \qquad \text{(III.12)}$$

Where: $M$is the filter window size, $x(k)$is the input signal, $y(k)$is the filtered signal.

Low-pass filtering.

Savitsky–Golay filtering.

Wavelet-based denoising.

Noise reduction is particularly important when extracting statistical and derived features from current and voltage signals.

III.2.4.4 Dimensionality Reduction Using PCA:

As the number of extracted features increases, redundant and highly correlated information may be introduced into the dataset. High-dimensional feature spaces can increase computational complexity and lead to overfitting. Principal Component Analysis (PCA) is a widely used dimensionality reduction technique that transforms correlated variables into a smaller set of uncorrelated variables called principal components. The PCA transformation can be expressed as:

$$Z = W^T X \tag{III.13}$$

where:

- X is the original feature matrix.
- W is the projection matrix formed by the eigenvectors of the covariance matrix.
- Z is the transformed feature matrix in the principal component space.

The principal components are ranked according to the amount of variance they explain. The first principal component captures the maximum variance, while subsequent components capture progressively smaller amounts of information. Consequently, PCA serves as an effective preprocessing tool for AI-based fault diagnosis systems, particularly when a large number of electrical, statistical, and derived features are extracted from PV measurements.

III.3 Artificial Neural Network (ANN) Approach:

The ANN is a computational model inspired by the biological neural network of the human brain, composed of interconnected processing units called neurons. ANNs learn complex relationships between inputs and outputs through a training process that adjusts connection weights to minimize prediction errors. In photovoltaic fault diagnosis, ANNs can automatically recognize patterns in measured data and classify different fault conditions with high accuracy.

III.3.1 ANN Architecture:

The architecture of ANN consists of three main layers: an input layer, one or more hidden layers, and an output layer. Illustrated in table III.

Input layer: extracted features, hidden layers: nonlinear mapping and output layer: fault classes.

Table III.3: Illustrated the architecture of ANN.

Layer	role	Design considerations
Input layer	Receives extracted features (e.g., PV I-V curve features, PCA components, statistical features)	Number of neurons = number of features (e.g., after PCA: 5–15 components)
Hidden layer	Nonlinear mapping to capture complex fault patterns	1–3 hidden layers; 8–64 neurons each; activation = Re LU/tanh
Output layer	Fault class prediction	Neurons = number of fault classes; activation = softMax for multi-class s

III.3.2 Training and Optimisation:

The training and optimization process is a fundamental stage in the development of photovoltaic (PV) fault diagnosis models. It ensures that the neural network learns meaningful relationships between input features and output fault classes by minimizing the prediction error. This process is typically performed using the backpropagation algorithm combined with gradient-based optimizations techniques [59]. Backpropagation is the core learning mechanism used in neural networks to update model weights. It works by propagating the error from the output layer back through the hidden layers and adjusting weights in order to minimize the loss function. The weight update rule is defined as:

$$W_{t+1} = W_t - \eta \frac{\partial L}{\partial W} \quad \text{(III.13)}$$

where:

- W_t is the current weight,

- $W_{t+1}$is the updated weight,
- $\eta$is the learning rate,
- $L$is the loss function.

This iterative process continues until the model converges to an optimal solution with minimal error [59].

III.3.2.1 Learning Rate:

The learning rate controls how much the model weights are updated during training. It plays a critical role in convergence:

- A high learning rate may cause instability and divergence.
- A low learning rate may lead to slow convergence.

Typical values range from $10^{-4}$to 10^{-2}, depending on the dataset and model architecture.

III.3.2.2 Epochs:

an epoch represents one complete pass through the entire training dataset. Increasing the number of epochs allows the model to learn more complex patterns; however, too many epochs may lead to overfitting. Early stopping is often used to prevent overfitting by stopping training when validation performance stops improving.

III.3.2.3Activation Functions:

Activation functions introduce non-linearity into the model, enabling it to learn complex relationships in PV fault data.

A. **Re LU (Rectified Linear Unit):** Re LU is widely used in hidden layers due to its simplicity and ability to reduce vanishing gradient problems. it is defined as:

$$\mathrm{f(x)} = \max(0, \mathrm{x}) \quad \text{(III.15)}$$

B. **Sigmoid Function:** s a nonlinear activation function commonly used in Artificial Neural Networks (ANNs) to transform input values into outputs ranging between 0 and 1.it is defined as:

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad \text{(III.16)}$$

C. **SoftMax Function:** SoftMax is used in the output layer for multi-class PV fault classification, converting outputs into probability distributions. it is defined as:

$$P(y_i) = \frac{e^{z_i}}{\sum_j e^{z_j}} \qquad \text{(III.17)}$$

D. **Loss function cross-entropy:** Cross-entropy loss is commonly used for classification problems because it measures the difference between predicted probabilities and actual class labels. For multi-class classification, it is defined as:

$$L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \qquad \text{(III.18)}$$

where:

- C is the number of classes,
- y_i is the true label,
- $\hat{y}_i$ is the predicted probability.

Minimizing cross-entropy ensures that the model assigns high probability to the correct fault class while penalizing incorrect predictions [59].

III.3.2.4 Performance Evaluation:

The performance of the photovoltaic (PV) fault diagnosis models is evaluated using standard classification metrics. These metrics provide a comprehensive assessment of the model's ability to correctly identify and distinguish between normal and faulty operating conditions. The most commonly used evaluation indicators include accuracy, precision, recall, F1-score, confusion matrix, and ROC curve analysis [59, 61]

a. **Accuracy:** measures the overall proportion of correctly classified samples among all predictions:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \qquad \text{(III.19)}$$

Where:

- TP: True Positives
- TN: True Negatives

- FP: False Positives
- FN: False Negatives

Accuracy provides a general indication of model performance but may be insufficient when dealing with imbalanced datasets.

b. **Precision**: evaluates the proportion of correctly identified fault cases among all predicted faults:

$$Precision = \frac{TP}{TP+FP} \quad \text{(III.20)}$$

High precision indicates a low false alarm rate, which is important in PV systems to avoid unnecessary maintenance actions.

c. **Recall (Sensitivity)** measures the ability of the model to correctly detect actual faults:

$$Recall = \frac{TP}{TP+FN} \quad \text{(III.21)}$$

High recall ensures that most fault conditions are successfully detected.

d. **F1-score** combines precision and recall into a single metric:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad \text{(III.22)}$$

It provides a balanced evaluation, especially when class distributions are uneven.

III.3.2.4 Confusion Matrix:

The confusion matrix is a detailed performance evaluation tool that summarizes the classification results by comparing predicted labels with actual labels. Each row represents the actual class, while each column represents the predicted class. It allows visualization of correct classifications and misclassifications across different fault categories. A typical confusion matrix helps in identifying:

- Which fault types are frequently misclassified.
- The overall classification strength of the model.
- Class imbalance effects on performance.

Confusion matrices are widely used in PV fault diagnosis because they provide detailed insight into model behavior beyond single-value metrics [23].

III.3.2.5ROC Curve

The Receiver Operating Characteristic (ROC) curve is used to evaluate the diagnostic ability of the classification model across different threshold values. It plots: True Positive Rate (TPR) against, False Positive Rate (FPR) .

$$TPR = \frac{TP}{TP+FN}, FPR = \frac{FP}{FP+TN} \tag{23.III}$$

Where: The Area Under the Curve (AUC) is a key indicator of performance: AUC = 1 → perfect classification, AUC = 0.5 → random classification. A higher AUC value indicates better model discrimination capability between faulty and normal PV conditions [23].

III.4 Convolutional Neural Network (CNN) Approach

III.4.1 Motivation for CNN:

Convolutional Neural Networks (CNNs) have emerged as one of the most effective deep learning techniques for pattern recognition and fault diagnosis due to their ability to automatically extract meaningful features from complex data. In photovoltaic (PV) fault diagnosis, CNNs are particularly attractive because they can analyze the nonlinear characteristics of PV electrical behavior without requiring extensive manual feature engineering [65].

III.4.2Treating I–V Curves as Images:

The current–voltage (I–V) characteristic curve of a photovoltaic module contains valuable information about its operating condition and health status. Different fault types, such as partial shading, open-circuit faults, short-circuit faults, hotspots, and degradation, produce distinctive distortions in the shape of the I–V curve. By converting I–V curves into graphical representations or image-like matrices, CNNs can analyze these patterns similarly to image recognition tasks [56].

Unlike traditional machine learning methods that rely on manually extracted electrical features, CNNs can directly process I–V curve images and learn fault-specific characteristics from the raw data. This approach enables the model to capture subtle variations and complex fault signatures that may be difficult to identify using conventional techniques.

III.4.3Automatic Feature Extraction:

One of the main advantages of CNNs is their ability to perform automatic feature extraction. Traditional fault diagnosis methods require the manual selection of relevant features such as current, voltage, power, fill factor, and statistical indicators. The effectiveness of these approaches often depends on expert knowledge and the quality of the selected features [61].

III.4.4 CNN Architecture:

The convolutional neural network (CNN) architecture is designed to automatically extract relevant features from input data and perform fault classification with high accuracy. Unlike conventional Artificial Neural Networks (ANNs), CNNs employ convolution and pooling operations to learn hierarchical feature representations, making them particularly effective for recognizing complex patterns associated with photovoltaic (PV) faults [61]. The CNN architecture consists of four main components, the convolutional layers, pooling layers, fully connected layers, and an output layer for fault classification.

III.4.4.1The convolutional layers:

Convolution layers constitute the core component of the CNN. Their primary function is to extract local features from the input data through the application of learnable filters (kernels). During training, these filters automatically learn important fault-related characteristics such as variations in electrical signals, I–V curve patterns, or thermal image features. The convolution operation is defined as:

$$Y(i,j) = \sum_m \sum_n X(i-m, j-n)K(m,n) \tag{III.24}$$

where: X is the input feature map, K is the convolution kernel, Y is the resulting feature map.

III.4.4.2The pooling layers:

Pooling layers are used to reduce the spatial dimensions of the feature maps while retaining the most significant information. This operation decreases computational complexity and helps prevent overfitting. The most commonly used pooling technique is Max Pooling:

$$Y = \max(X) \tag{III.25}$$

where the maximum value within a predefined pooling window is selected as the output.

III.4.4.3The fully connected layers:

After feature extraction through convolution and pooling operations, the resulting feature maps are flattened into a one-dimensional vector and passed to fully connected layers. The output of a fully connected neuron is computed as:

$$z = \sum_{i=1}^{n} w_i\, x_i + b \tag{III.26}$$

Where: $x_i$are the input features, $w_i$are the connection weights, b is the bias term, z is the neuron output.

III.4.4.4The output layer for fault classification:

The final layer of the CNN performs the fault classification task. The number of output neurons corresponds to the number of classes considered in the diagnostic system, such as: Normal operation, Partial shading fault, Open-circuit fault, Short-circuit fault, Degradation fault, Hotspot fault. For multiclass classification problems, the SoftMax activation function is commonly employed:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}} \tag{III.27}$$

where: $P(y_i)$is the probability of class i, $z_i$is the output score of class i, $C$is the total number of classes. The SoftMax layer converts the network outputs into probability values, and the class with the highest probability is selected as the predicted fault category [61].

III.4.5 Training and Hyperparameter Tuning:

The performance of a Convolutional Neural Network (CNN) largely depends on the selection of appropriate training parameters and hyperparameters. Hyperparameter tuning aims to optimize the learning process, improve classification accuracy, and enhance the generalization capability of the model. In photovoltaic (PV) fault diagnosis, proper tuning is essential to ensure reliable fault detection while avoiding overfitting and excessive computational costs [61].

III.4.5.1 Batch Size:

The batch size determines the number of training samples processed before updating the network weights. Smaller batch sizes generally improve generalization and reduce memory requirements, whereas larger batch sizes accelerate training but may lead to poorer convergence. Typical batch sizes used in CNN-based PV fault diagnosis range from 16 to 128 samples. The optimal value is selected experimentally according to dataset size and available computational resources [61].

III.4.5.2 Number of Epochs:

An epoch represents one complete pass through the entire training dataset. Increasing the number of epochs allows the CNN to learn more complex feature representations; however, excessive training may result in overfitting. The training process is repeated for a predefined number of epochs:

$$\text{Epochs} = N_{iter} \tag{III.28}$$

Where: N_{iter}denotes the total number of training iterations over the dataset. In practice, training is often stopped when the validation loss no longer decreases, using an early-stopping strategy to prevent overfitted [59].

III.4.5.3 Learning Rate:

The learning rate controls the magnitude of weight updates during the optimization process. It is one of the most important hyperparameters because it directly affects convergence speed and model stability. The weight update rule using gradient descent is:

$$W_{t+1} = W_t - \eta \nabla L(W_t) \tag{III.29}$$

where: /

- $W_t$is the current weight vector,

- η is the learning rate,
- $\nabla L(W_t)$is the gradient of the loss function,
- $W_{t+1}$is the updated weight vector.

A learning rate that is too large may cause unstable training, while a very small learning rate can result in slow convergence. Typical values range from $10^{-4}$to 10^{-2}, depending on the optimizer and dataset characteristics [62].

III.4.6 Regularization Methods:

To improve model generalization and reduce overfitting, regularization techniques are incorporated into the CNN training process. These methods help the model learn meaningful fault-related features rather than memorizing the training data.

III.4.6.1Dropout:

Dropout is a widely used regularization technique that randomly deactivates a fraction of neurons during training. This prevents excessive dependence on specific neurons and encourages the network to learn more robust feature representations [46].

If $p$represents the dropout probability, each neuron is retained with probability:

$$P(\text{retain}) = 1 - p \tag{30.III}$$

Typical dropout values range between 0.2 and 0.5. In PV fault diagnosis applications, dropout improves robustness and reduces the risk of overfitting when training datasets are limited.

III.4.6.2 L2 Regularization (Weight Decay):

L2 regularization penalizes large weight values by adding a penalty term to the loss function:

$$L_{reg} = L + \lambda \sum_{i=1}^{n} w_i^2 \tag{31.III}$$

where:

- $L$is the original loss function,

- $\lambda$is the regularization coefficient,
- $w_i$represents the network weights.

L2 regularization encourages smaller weight magnitudes, resulting in smoother decision boundaries and improved generalization performance. This technique is particularly beneficial in PV fault diagnosis models where noisy measurements and environmental variations may otherwise lead to overfitting [59].

III.5 Performance Metrics:

To assess the effectiveness of PCA-based classifiers, their performance is compared with that of the Artificial Neural Network (ANN) model developed for photovoltaic (PV) fault diagnosis. ANN is a powerful machine learning technique capable of learning complex nonlinear relationships between input features and fault classes. Unlike PCA-based methods, which rely on dimensionality reduction followed by classification, ANN directly learns the mapping between extracted features and fault conditions through a network of interconnected neurons [47]. The comparison focuses on detection accuracy, robustness, computational complexity, and suitability for real-time implementation. ANN generally achieves higher classification accuracy because of its ability to model nonlinear patterns and interactions among electrical, statistical, and derived features. However, ANN requires a larger training dataset and greater computational resources than PCA-based approaches [59].

III.5.1 Evaluation of Detection Accuracy and Robustness:

The performance of the classifiers is evaluated using standard classification metrics such as accuracy, precision, recall, and robustness against noise and operating condition variations. Detection accuracy measures the ability of the model to correctly identify fault conditions, while robustness reflects its capability to maintain reliable performance under varying environmental conditions and noisy measurements.

III.5.2 Classification Accuracy:

Classification accuracy is a performance metric used to evaluate a classification model by measuring the proportion of correctly classified instances among all instances in the dataset. calculated as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \qquad \text{(III.32)}$$

where:

- TP= True Positives,
- TN= True Negatives,
- FP= False Positives,
- FN= False Negatives.

Higher accuracy indicates better fault identification capability.

III.5.3 Comparative Performance:

The comparison of the performance of the AI based classification is illustrated in the table III.4.

Table III.4: Illustrated AI comparison of classification.

Criterion	**ANN**	**PCA-Based Classifier (KNN/SVM)**
Detection Accuracy	High (typically 90–98%)	Moderate to High (85–96%)
Robustness to Nonlinear Faults	Excellent	Good
Computational Complexity	Moderate to High	Low
Training Time	Longer	Shorter
Real-Time Deployment	Good	Excellent
Large Dataset Handling	Good	Excellent
Memory Requirement	Higher	Lower

The results indicate that ANN generally provides superior fault detection accuracy due to its nonlinear learning capability. It is particularly effective when complex fault signatures and interactions between features are present. However, PCA-based classifiers offer faster training and lower computational requirements because dimensionality reduction eliminates redundant information before classification [53].

III.6 PCA-Based Classifier Approach:

The PCA-based classifier approach first applies principal component analysis (PCA) to reduce the dimensionality of the extracted features while preserving the most important information.

III.6.1 Dimensionality Reduction Using PCA:

III.6.1.1 Retaining Maximum Variance:

The primary objective of PCA is to retain the maximum possible variance contained in the original data using a reduced number of variables. The first principal component captures the largest variance in the dataset, while each subsequent component captures the maximum remaining variance subject to being orthogonal to the previous components [7]. The amount of variance explained by each principal component is determined by its corresponding eigenvalue λ_i. The cumulative explained variance ratio is given by:

$$\text{Explained Variance Ratio} = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \tag{III.33}$$

Where:

- $\lambda_i$is the eigenvalue of the i^{th}principal component,
- $k$is the number of retained principal components,
- $n$is the total number of original features.

In practical applications, enough principal components are retained to preserve approximately 90–95% of the total variance. This approach significantly reduces computational complexity while maintaining most of the information required for accurate fault classification. In photovoltaic fault diagnosis, PCA helps eliminate redundant features, improve classifier efficiency, and reduce the risk of overfitting, particularly when dealing with large datasets containing electrical, statistical, and derived features [47].

III.6.2 Classifier Implementation:

After dimensionality reduction using Principal Component Analysis (PCA), the transformed feature set is used to train and evaluate machine learning classifiers for photovoltaic (PV) fault diagnosis. In this study, two widely used classification algorithms, namely K-Nearest Neighbour (KNN) and Support Vector Machine (SVM), are implemented to identify and classify different PV fault conditions. These classifiers are selected because of their effectiveness in handling multidimensional datasets and their proven performance in fault diagnosis applications [60].

III.6.2.1 K-Nearest Neighbors (KNN):

K-Nearest neighbors (KNN) is a supervised learning algorithm that classifies a new sample based on the classes of its nearest neighbors in the feature space. The classification decision is made according to the majority vote among the Kclosest training samples [23].

The Euclidean distance is commonly used to measure similarity between samples:

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{n} (x_{ik} - x_{jk})^2} \qquad \text{(III.34)}$$

where:

- $d(x_i, x_j)$is the distance between two samples,
- $n$is the number of features,
- $x_{ik}$and $x_{jk}$are feature values of the compared samples.

KNN is simple to implement, requires no explicit training phase, and performs well for datasets with clearly separated fault classes. However, its computational cost increases with dataset size because distances must be calculated for each new observation.

III.6.2.2 Support Vector Machine (SVM):

Support Vector Machine (SVM) is a powerful supervised classification technique that constructs an optimal decision boundary, known as a hyperplane, to maximize the separation between classes [60].

The linear SVM decision function is expressed as:

$$\mathrm{f(x)} = \mathrm{w^T x + b} \qquad \text{(III.35)}$$

where:

- w is the weight vector,
- x is the input feature vector,
- b is the bias term.

For nonlinear fault classification problems, kernel functions such as the Radial Basis Function (RBF) can be employed to project data into higher-dimensional spaces where classes become more separable. SVM is known for its strong generalization capability and effectiveness in high-dimensional datasets, making it suitable for PV fault diagnosis after PCA feature reduction.

III.6.3 Comparison of Classifier Performance:

The performance of KNN and SVM classifiers is evaluated using the PCA-transformed feature dataset. Classification accuracy, computational complexity, and robustness are considered as the primary evaluation criteria.

Table III.4: illustrated the comparison classification of AI performance.

Classifier	**Typical Accuracy (%)**	**Advantages**	**Limitations**
KNN	85–95	Simple implementation; no training phase; effective for small and medium datasets; easy to understand.	High computational cost during testing; sensitive to noise and choice of (K); performance decreases with large datasets.
SVM	90–98	High classification accuracy; strong generalization capability; effective in high-dimensional feature spaces; robust to overfitting.	Requires parameter tuning; higher training complexity; kernel selection influences performance.

III.6.3.1Comparative Analysis of AI Models:

Artificial Intelligence (AI) techniques have become increasingly important for photovoltaic (PV) fault diagnosis due to their capability to identify complex patterns and classify fault conditions with high accuracy. Among the most widely used approaches are Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and Principal Component Analysis (PCA)-based classifiers. Each technique offers distinct advantages and limitations depending on the characteristics of the dataset, computational resources, and application requirements. Illustrated the comparative in this table III.5.

Table III.5: comparison of the analysis of AI models.

Model	**Typical Accuracy**	**Advantages**	**Limitations**
ANN	90–98%	Simple implementation, fast training, suitable for electrical parameters such as voltage, current, and power	Performance depends on feature quality and network architecture; may struggle with highly complex patterns
CNN	95–99%	Automatic feature extraction, excellent fault classification capability, robust to nonlinear relationships and complex fault signatures	Higher computational cost, requires larger datasets and more training time
PCA-Based Classifiers	85–95%	Reduces dimensionality, decreases computational burden, improves processing speed, suitable for large datasets	Possible loss of useful information during dimensionality reduction; diagnostic accuracy may be lower than deep learning models

III.6.4 Observations:

The comparative analysis of the Artificial Intelligence (AI) models highlights distinct strengths and limitations associated with each approach. The obtained results indicate that the effectiveness of a diagnostic model depends not only on its classification accuracy but also on the nature of the photovoltaic (PV) data, computational requirements, and deployment constraints.

III.6.4.1 CNN Excels for Complex Fault Signatures:

The Convolutional Neural Network (CNN) achieved the highest diagnostic performance when dealing with complex fault patterns and highly nonlinear PV operating conditions. CNN architectures are capable of automatically extracting hierarchical features from input data, allowing them to identify subtle variations associated with partial shading, hotspot formation, module degradation, and multiple simultaneous faults.

35This capability makes CNNs particularly suitable for analyzing I–V curve images, thermal images, and large multidimensional datasets. Previous studies have reported that CNN-based models consistently outperform conventional machine learning methods in complex PV fault classification tasks due to their superior feature-learning capabilities [56, 57].

III.6.4.2 ANN Is Effective for Tabular Electrical Features:

Artificial Neural Networks (ANNs) demonstrated strong performance when trained using structured numerical features extracted from PV measurements, including current (I), voltage (V), power (P), fill factor (FF), and statistical indicators.

Since these features already contain relevant diagnostic information, ANN models can efficiently learn the relationships between input variables and fault categories without requiring sophisticated feature extraction procedures. In addition, ANNs offer relatively low computational complexity, fast inference speed, and straightforward implementation, making them highly suitable for real-time PV monitoring systems and embedded diagnostic applications [47].

III.6.4.3 PCA-Based Methods Are Computationally Efficient for Large Datasets:

Principal Component Analysis (PCA)-based methods proved particularly effective in reducing computational complexity when handling large PV datasets. By transforming the original feature space into a reduced set of principal components, PCA removes redundant and highly correlated information while preserving most of the variance contained in the data. This dimensionality reduction significantly decreases training and processing time, thereby improving the efficiency of subsequent classifiers such as Support Vector Machines (SVMs) and K-Nearest Neighbors (KNNs) [6]. Consequently, PCA-based approaches are especially attractive for large-scale PV monitoring systems and resource-constrained environments where computational efficiency is a critical requirement.

III.7 Deployment Considerations:

Successfully integrating an AI-driven photoelectric fault detection (PV) system necessitates a strategic approach to implementation to guarantee reliable functioning under real-world conditions. In addition to achieving high classification accuracy, the diagnostic system must support real-time monitoring, seamless integration with existing PV monitoring infrastructures, and efficient hardware implementation. These factors are crucial for enabling ongoing fault detection, minimizing outages, and enhancing the overall efficiency and dependability of photovoltaic setups [47].

III.7.1 Real-time monitoring:

It is crucial for modern photovoltaic systems, especially in large solar farms, as it aids in rapid fault detection, minimizing energy losses and maintenance costs. An AI-based diagnostic framework uses electrical and statistical features from continuously measured PV signals like current, voltage, and power for real-time fault classification with trained machine learning models. The deployment feasibility relies on computational complexity, data acquisition frequency, and communication latency. Among the methods, Artificial Neural Networks (ANNs) are favored for their quick inference times, while Convolutional Neural Networks (CNNs) offer higher accuracy but demand more computational resources. PCA-based classifiers help reduce feature dimensionality, enhancing suitability for real-time applications due to their processing speed. [47, 49]. The advantages of real- time monitoring as:

1. Continuous acquisition of PV electrical measurements, Feature extraction and preprocessing.
2. Dimensionality reduction using PCA, AI-based fault classification.
3. Alarm generation and maintenance notification

II.7.2 Integration with PV monitoring systems:

Modern photovoltaic installations are commonly equipped with Supervisory Control and Data Acquisition (SCADA) systems, data loggers, smart inverters, and Internet of Things (IoT) sensors. The proposed fault diagnosis framework can be integrated into these existing monitoring infrastructures without requiring major hardware modifications [50]. The AI diagnostic module can receive data from:

- Solar irradiance sensors.
- Temperature sensors.
- Voltage sensors.
- Current sensors.
- Smart inverters.
- Energy meters.
- Data acquisition systems.

The integration process typically involves:

- Data collection from monitoring devices.
- Signal preprocessing and feature extraction.
- AI-based fault diagnosis.
- Storage of diagnostic results in databases.
- Visualization through monitoring dashboards, Automated maintenance alerts.

Such integration allows the fault diagnosis system to operate as an additional intelligent layer within the PV monitoring architecture. Diagnostic results can be displayed through web-based interfaces, mobile applications, or SCADA dashboards, enabling operators to monitor system health remotely and respond promptly to detected anomalies [51]. Furthermore, cloud-based deployment enables centralized monitoring of multiple PV installations simultaneously, making the solution scalable for utility-scale solar farms and distributed PV networks**.**

III.7.3 Hardware implementation options (edge devices, IoT platforms):

The proposed AI-based fault diagnosis system can be deployed using edge computing devices and IoT platforms to enable real-time monitoring and fault detection in photovoltaic (PV) systems. Edge devices such as Raspberry Pi and NVIDIA Jetson Nano can perform local data processing and fault classification, reduce latency and minimize dependence on cloud services. These platforms are particularly suitable for implementing ANN and PCA-based diagnostic models [47].

IoT platforms facilitate data collection, transmission, and remote monitoring through communication technologies such as Wi-Fi, LoRaWAN, ZigBee, and 4G/5G networks. Sensor measurements, including voltage, current, irradiance, and temperature, can be transmitted to monitoring servers for visualization and maintenance management [51].

A hybrid edge–IoT architecture combines local fault diagnosis with cloud-based storage and analytics, providing a scalable, cost-effective, and reliable solution for intelligent PV monitoring systems [58].

III.8 Conclusion:

In the end, artificial intelligence (AI) has shown a lot of promise for making fault analysis in photovoltaic (PV) systems better by making detection quick, accurate, and automatic. AI-based methods are better than traditional ways of diagnosing problems because they can quickly look at a lot of electrical data and identify complicated fault patterns in various environments.

Convolutional neural networks (CNNs) were the best at finding complex and nonlinear fault signals out of all the methods used. This is because they are very good at extracting features and recognizing patterns. They are great for advanced fault classification jobs because they can automatically learn features that make them different from PV datasets.

Artificial Neural Networks (ANNs), on the other hand, are an easier and less computationally intensive option that can still accurately classify things. Because ANN models are easy to set up and use less computing power, they can be used for real-time monitoring and embedded diagnostic systems. Also, Principal Component Analysis (PCA)-based algorithms are a good way to reduce the number of dimensions in a dataset while keeping the most important data.

PCA speeds up computations and makes the classification process easier by getting rid of unnecessary features. This is especially helpful when working with big modelling datasets. Overall, this chapter makes a strong link between creating data through simulations and using clever diagnostics in real life. The framework shows how simulated PV fault data can be processed, analysed, and combined with AI methods to create fault diagnosis systems for current photovoltaic uses that are reliable and effective.

.

Chapter VI:

Results, Validation, and Discussion

IV.1 Introduction:

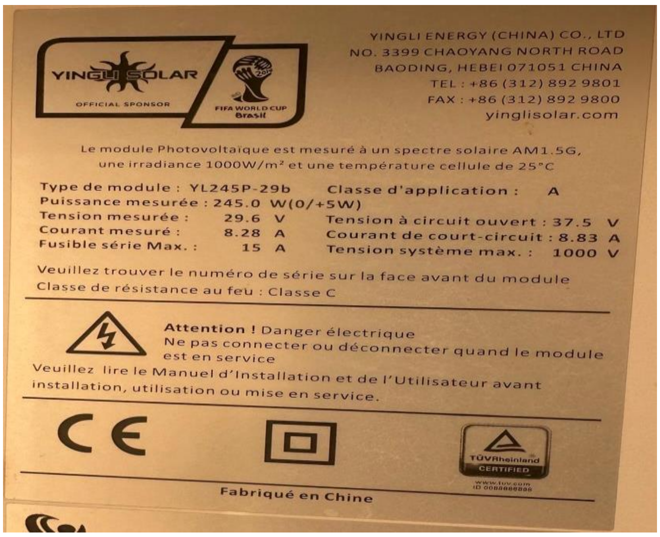
This chapter presents the simulation results and performance evaluation of the proposed PCA–ANN fault diagnosis system for photovoltaic installations. The study investigates the electrical behavior of the photovoltaic generator under normal and faulty operating conditions using MATLAB/Simulink. Different fault scenarios, including partial shading, series resistance increase (Rs), shunt resistance decrease (Rsh), and bypass diode failure, were introduced into the system to evaluate their impact on the electrical characteristics of the PV array.

The extracted electrical parameters were processed using Principal Component Analysis (PCA) to reduce data dimensionality and eliminate redundant information. The reduced feature set was subsequently used as input to an artificial neural network (ANN) classifier responsible for fault identification and classification. The obtained results are analyzed from electrical, statistical, and diagnostic perspectives.

IV.2 The characteristically panel using in simulation:

Represented the characteristics of PV panel using in Simulation in the table IV.1 is illustrated the parameter of panel of the PV system.

Table IV.1: Characteristic PV Panel.

<table>
<tr><td>Parameter PV panel:</td><td rowspan="8">
</td></tr>
<tr><td>Type of module: YL245P-29b
CLASSE A</td></tr>
<tr><td>P:245,0W (0/+5W)</td></tr>
<tr><td>Isc:8.83A</td></tr>
<tr><td>Voc:37.5V</td></tr>
<tr><td>Ipv: 8.28</td></tr>
<tr><td>Vpv:29.6</td></tr>
<tr><td>System Voltage 1000V</td></tr>
</table>

IV.3 Simulation Environment and Experimental Protocol:

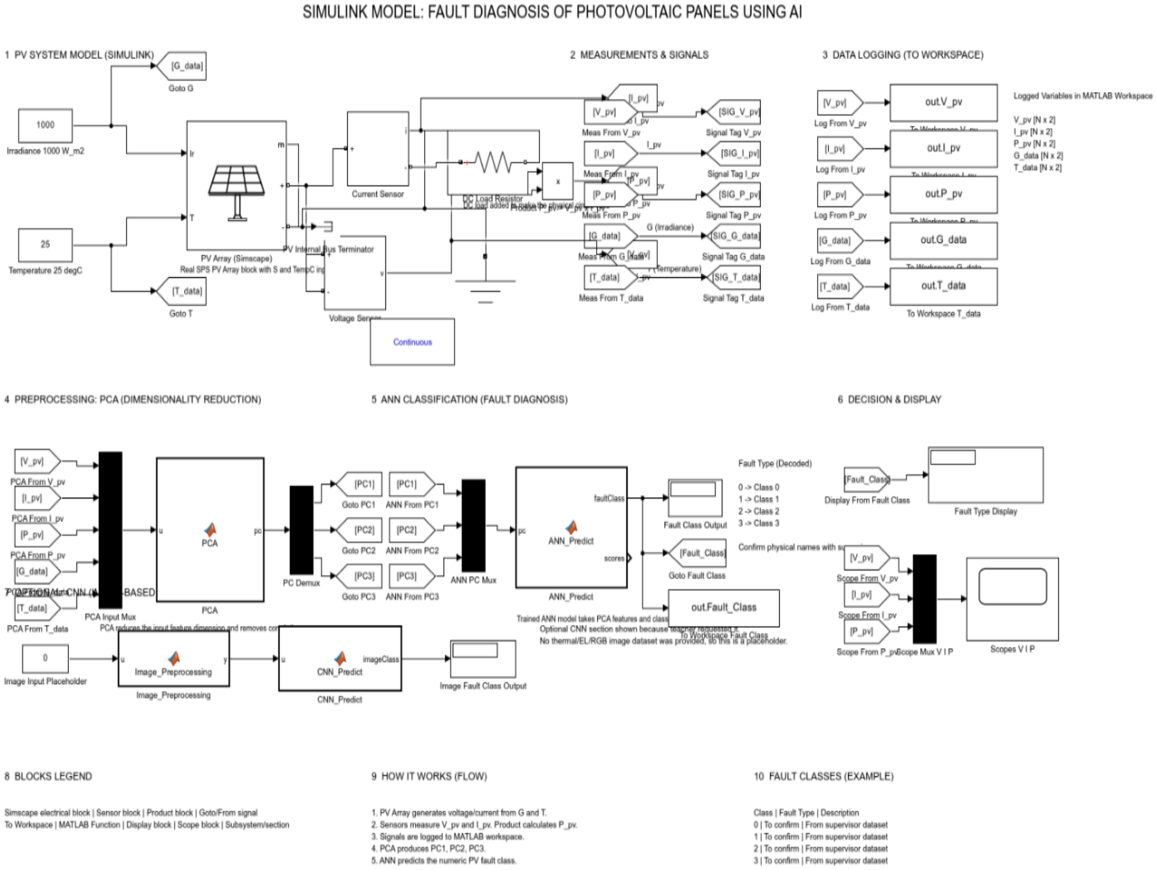


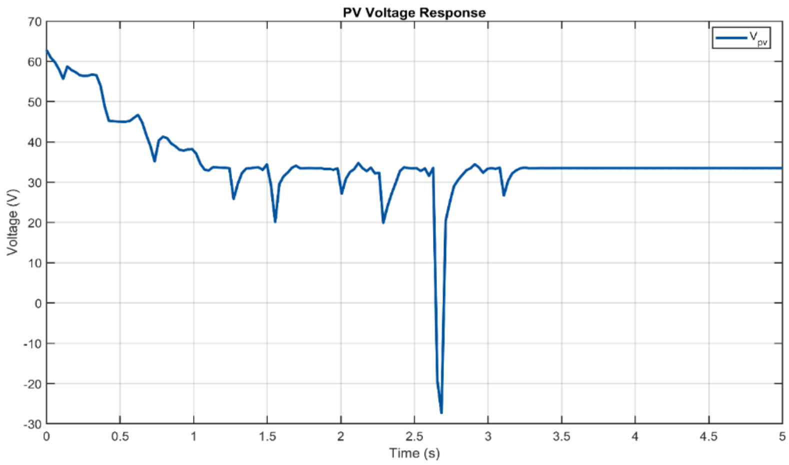
Figure IV.1: Diagram of Simulink model of fault diagnosis PV panel using AI.

IV.4 Analysis of the Electrical Performance of the Photovoltaic System:

The analysis of the electrical performance of a photovoltaic (PV) system is essential for evaluating its efficiency, reliability, and energy production capability under different operating conditions. Key electrical parameters such as voltage, current, and output power are examined to assess the system's behavior in both normal and faulty states.

IV.4.1Analysis of the Evolution of Photovoltaic Voltage:

In the figure IV.2 represented the result of voltage response -0.0s – 3.2s (Transient/Perturbation). The PV system undergoes MPPT tracking and severe environmental or diode-switching disturbances, highlighted by a massive transient voltage drop to -27V at 2.65 seconds. 3.2s – 5.0s (Steady-State): The system successfully stabilizes and flatlines at an optimal operating voltage of 33.5V, indicating complete convergence of the tracker under steady conditions.



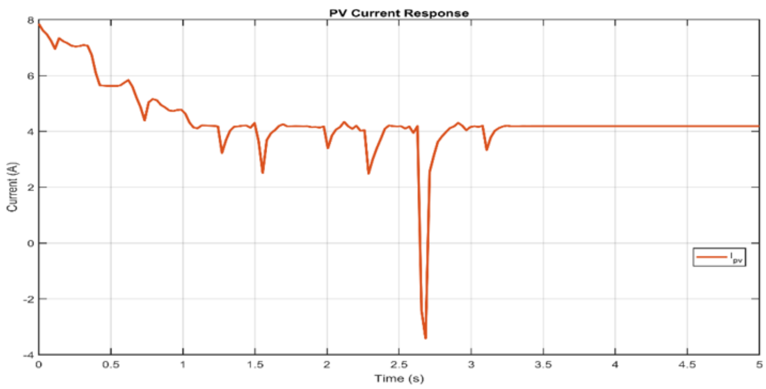
FigureIV.2: PV Voltage Response:

IV.4.2Analysis of the Evolution of Photovoltaic Current:

Extended Technical Analysis Initial Tracking (0.0s – 1.1s). The current starts near the short-circuit/open-circuit transition region at nearly 8A. The stepped descent directly mirrors the voltage plot's plateaus as the MPPT controller scans the PV characteristic curve to find the optimal operating point.

The Negative Current Peak (~2.65s): A negative current spike of -3.4A matches the massive negative voltage drop seen in your previous plot, see the figure IV.3.

This indicates a severe back-feeding event. In a physical solar array, this happens when a parallel string or the power converter's inductive elements force current backward through a temporarily shaded or faulted string, momentarily overwhelming the bypass and blocking diodes before recovering. In the System Characteristics are Because both the voltage Vpv= 33.5 V and current Ipv= 4.2A stabilize perfectly after 3.2 seconds, the steady-state power output of your simulated PV module/array settles at approximately.



FigureIV.3: PV Current Response

IV.4.3 Analysis of the Evolution of Photovoltaic Power:

The PV power decreases rapidly from approximately **490 W to 140 W**, exhibiting several sharp drops and oscillations that indicate the presence of disturbances such as partial shading or fault conditions affecting the photovoltaic array.

After **3.2 s**, the power stabilizes around **140 W**, showing that the system reaches a new steady-state operating point, although with a significant reduction in power output compared to normal conditions. see the figure IV.4.

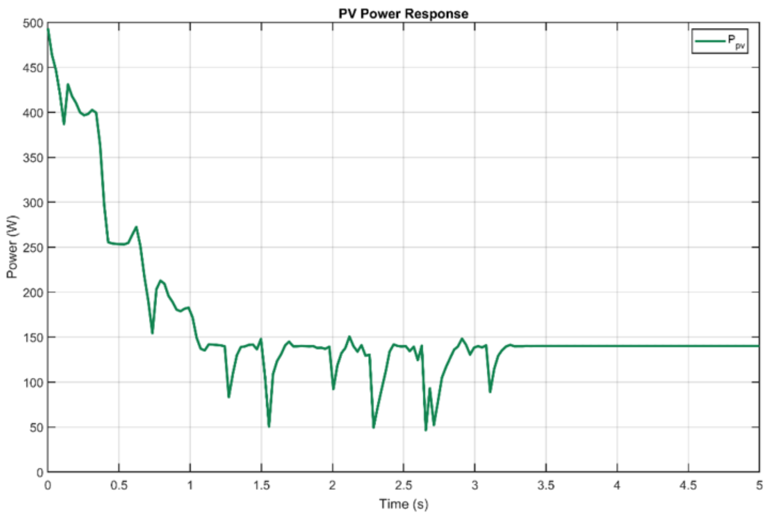


Figure IV.4: PV Power Response

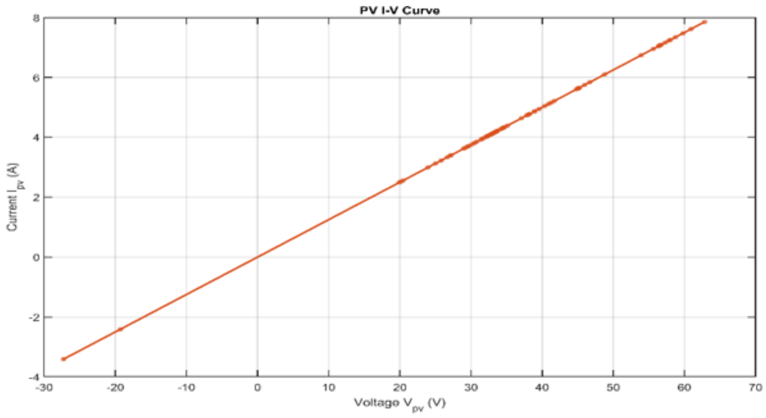
IV.4.4 Analysis of the Dynamic behaviour of the System in the Presence of Faults:

The dynamic responses of the PV system under fault conditions reveal significant disturbances in voltage, current, and power profiles. The PV voltage exhibits several sudden drops, including a severe voltage collapse around 2.7 s, indicating the occurrence of a critical fault. Similarly, the PV current experiences sharp fluctuations and a pronounced reduction at the same instant, leading to a substantial decrease in the generated power. Consequently, the PV power output shows the highest sensitivity to faults, with large transient losses and a reduced steady-state value, confirming the strong impact of faults on overall system performance, (See figures IV.2,3,4).

IV.5 Study of the Electrical Characteristics of the Photovoltaic Generator:

IV.5.1 Analysis of the Current-Voltage Characteristic (I–V):

Linear resistor characteristic the plot shows a completely straight, linear (I-V) relationship passing from positive values down through the origin into the third quadrant (reaching-27V, -3.4A), indicating that the system is operating like a passive resistor rather than a non-linear photovoltaic semiconductor. Absence of Diode Behavior. The curve entirely lacks the classic exponential knee region of a standard PV cell curve, demonstrating that under these simulation settings, the internal diode properties are bypassed, short-circuited, or overridden by a dominant resistive load or a severe line-to-line fault. Operating point concentration a dense cluster of data points is concentrated around 33.5V and 4.2A, highlighting where the system spends its stable steady-state phase after recover**.** See figure IV.5.



FigureIV.5: PV I-V Curve

IV.5.2 Analysis of the Power-Voltage Characteristic (P–V):

IV.5.2.1Parabolic Power Trajectory: The curve exhibits a prominent quadratic shape in the positive quadrant (P = V. I), directly tracking the power profile resulting from the linear relationship of your resistor-like circuit rather than the classic bell-shaped curve of an operational PV module.

IV.5.2.2Negative Voltage Power Spike: As the voltage sweeps deep into negative territory (down to 27 V, the power registers as a positive value sim 90W), showing the mathematical cross-product of negative current and negative voltage (V times -I = P) during the extreme transient fault event.

IV.5.2.3 Steady-State Convergence: The plot shows a high density of overlapping data points clustered tightly along the curve at the 33.5V mark, capturing where the system firmly anchors itself at its 140.7 W steady-state operating point once tracking stabilizes. See figureIV.6.

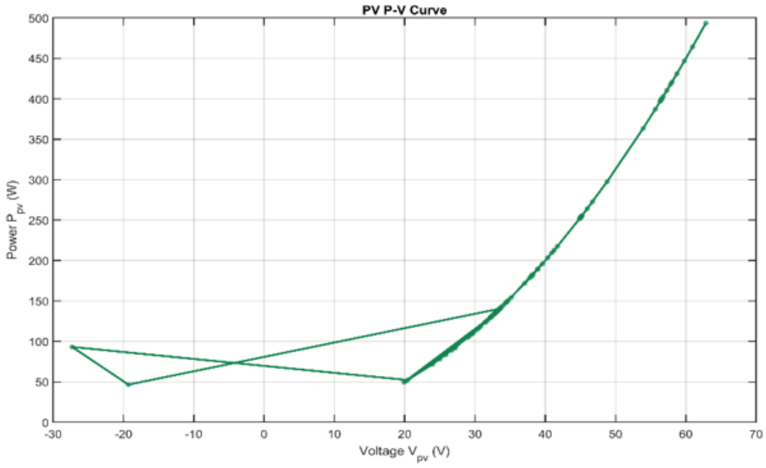


Figure IV.6: PV P-V Curve.

IV.6 Extraction and Analysis of Characteristics by Principal Component Analysis

IV.6.1 Analysis of the Extracted Principal Components:

Dynamic Feature Mapping (0.0s – 3.2s): The three principal components tightly track the system's transient states, capturing initial MPPT adjustments as a settling trend and isolating the critical fault at 2.65s through synchronized spikes across all channels (positive in PC1, negative in PC2 andPC3). See figure IV.7.

Diagnostic Steady-State (3.2s – 5.0s): After the perturbations clear, all three features lock into perfectly uniform, time-invariant flatlines (PC1≅22, PC2 ≅-18, PC3≅ -2.5), providing the ANN with a clean, stable signature for healthy steady-state classification.

IV.6.2 Influence of the Principal Components on Fault Detection:

a. **Signal Denoising and Feature Isolation:** PCA compresses noisy, multi-variable PV electrical data into highly distinct trajectories, ensuring that the initial tracking phases, minor shading dips, and steady-state conditions are clearly separated for the ANN.
b. **Fault Amplification:** At the **2.65s** mark, PCA magnifies the severe back-feeding fault by driving all three components to simultaneous, extreme transient

spikes, providing the neural network with an unmistakable, high-contrast signature for immediate fault detection.

c. **False-Alarm Reduction:** By flattening into rigid, uniform baselines (PC1 ≅22, PC2≅-18, PC3≅ -2.5) during the steady-state phase, the components eliminate normal operational noise, preventing the downstream classifier from triggering false alarms during healthy MPP tracking. Intelligent Fault Diagnosis Using Artificial Neural Networks.

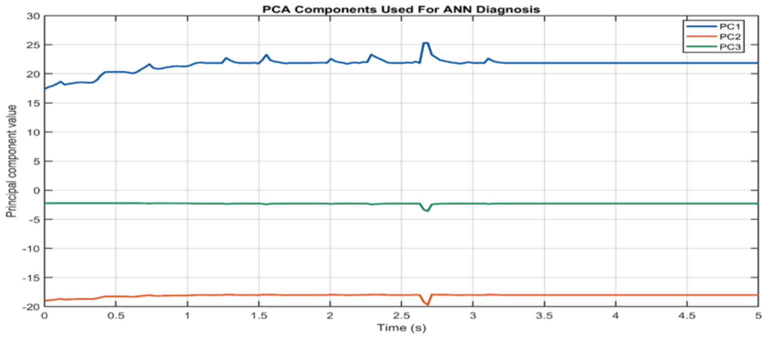
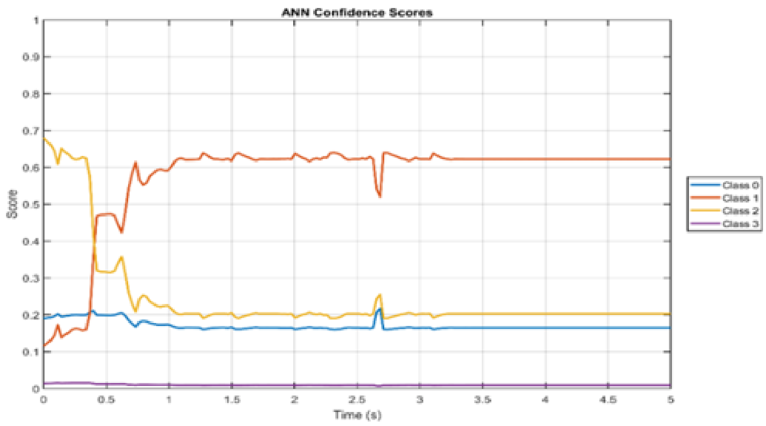


Figure IV.7: PCA Components Used for ANN Diagnosis

IV.6.3 Analysis of the ANN Classifier Confidence Scores:

The ANN confidence scores show a clear transition in classification during the transient period, where the dominant prediction shifts from **Class 2** to **Class 1** as the fault condition develops. After approximately **1 s**, **Class 1** stabilizes with the highest confidence score (≈0.62), indicating successful fault identification by the neural network. The low and nearly constant confidence levels of **Class 0** and **Class 3** confirm that these operating conditions are unlikely, demonstrating the robustness and reliability of the ANN classifier. See the figure IV.8.



FigureIV.8: ANN Confidence Scores.

IV.6.4 Identification of the Predicted Fault Class:

The graph displays a constant Fault Class ID of 1 predicted by the ANN over the entire time range from 0 to 5 seconds. In the temporal evolution of fault diagnosis is the graphs show an ANN's fault classification over time, with predictions remaining constant. Confidence scores for each class vary, indicating differing levels of prediction certainty. See figure IV.9.

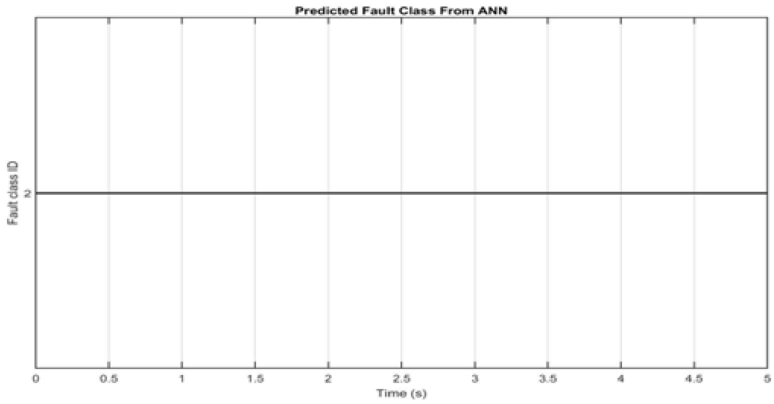


Figure IV.9: Predicted Fault Class From ANN

IV.7 Validation of the Diagnostic System Performance:

Tabel IV.2: illustrated validation and performance.

Model	Family	Accuracy	Macro F1	Weighted F1
Bagged Trees	Machine Learning	86.00%	0.8586	0.8586
LSTM Feature Sequence	Deep Learning	71.00%	0.7126	0.7126
CNN Feature Vector	Deep Learning	51.00%	0.5018	0.5018
KNN	Machine Learning	45.00%	0.4249	0.4249
Nearest Centroid Baseline	Baseline	33.00%	0.3267	0.3267
ANN Pattern Network	Deep Learning	31.00%	0.2367	0.2367
SVM ECOC	Machine Learning	26.00%	0.1234	0.1234

IV. 8 Global Dashboard and Summary of Results: In this section we have illustrates all the results for the simulation from the Simulink model of the fault diagnosis PV using AI on the figure forma. Illustrated in figure IV.10**.**

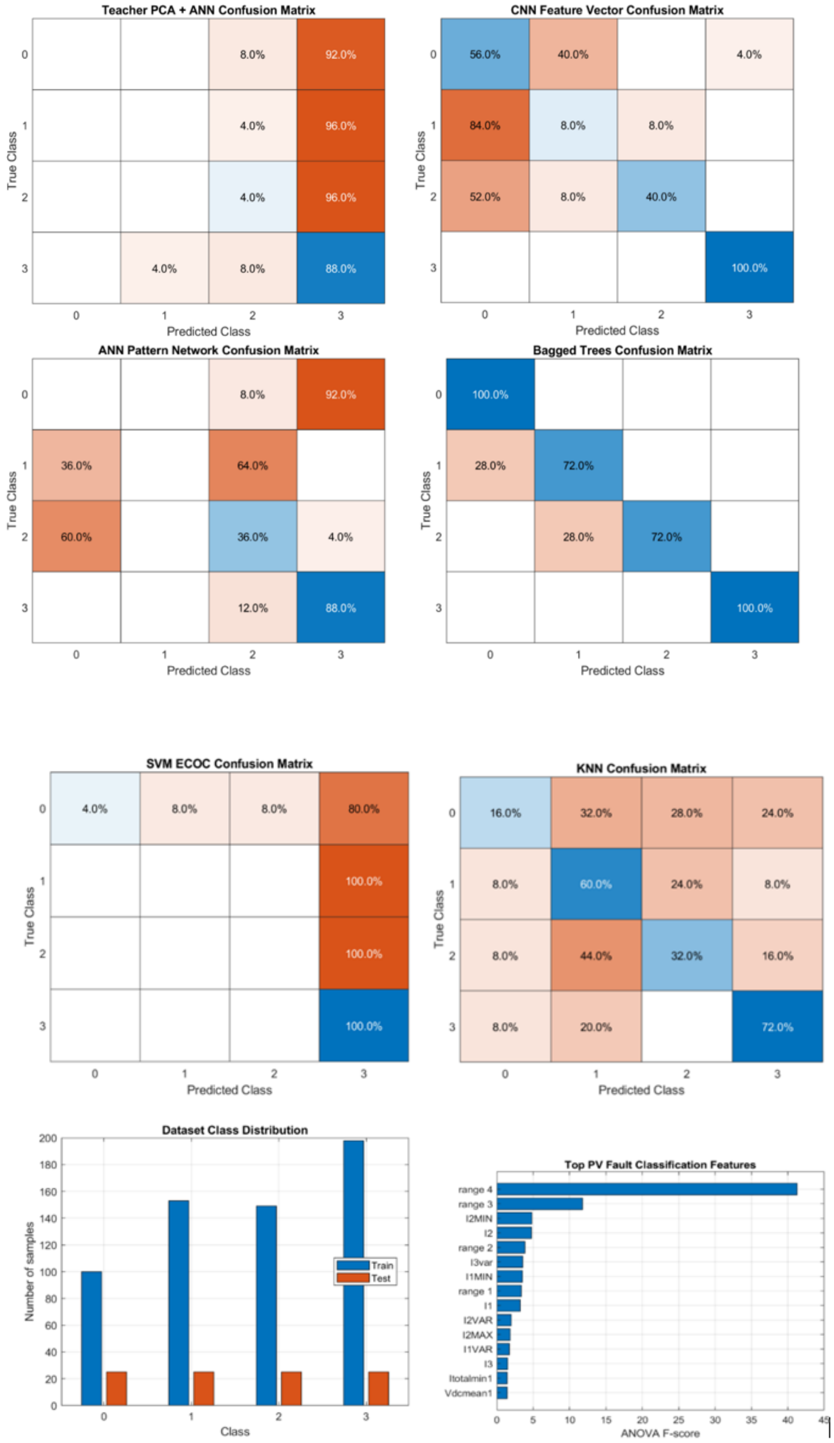


Figure IV.10: illustrated of the AI classification results.

IV.9 Discussion of the Electrical Performance of the Photovoltaic System:

The simulation results demonstrate that the electrical behavior of the photovoltaic (PV) system is highly sensitive to fault conditions. Under normal operation, the PV array produces stable voltage, current, and power outputs corresponding to the available solar irradiance and temperature conditions. However, the introduction of faults such as partial shading, open-circuit faults, and short-circuit faults leads to significant deviations in the electrical characteristics. The voltage and current profiles exhibit noticeable fluctuations and transient disturbances, which directly affect the power output. Among the monitored variables, power proved to be the most sensitive indicator of system degradation, showing substantial reductions during fault occurrences. These variations confirm that electrical parameters contain valuable information that can be exploited for automatic fault detection and classification.

IV. 9.1 Discussion of the PCA-ANN Model Performance:

The PCA-ANN diagnostic model demonstrated a strong capability to identify and classify PV system faults based on extracted electrical features. Principal Component Analysis (PCA) effectively reduced the dimensionality of the feature space while preserving the most significant information related to fault signatures. This preprocessing stage simplified the learning process and reduced computational complexity. The Artificial Neural Network (ANN) successfully learned the relationships between the principal components and the fault classes, achieving stable classification results and clear confidence score separation. The results indicate that the proposed PCA-ANN framework provides reliable fault diagnosis, rapid convergence, and good discrimination between different operating conditions, making it suitable for real-time PV monitoring applications.

IV.9.2 Limitations of the Proposed Approach:

Despite its promising performance, the proposed methodology presents several limitations. First, the diagnostic model was developed and validated primarily using simulated data generated in MATLAB/Simulink, which may not fully represent all uncertainties encountered in real PV installations. Second, only a limited number of fault categories were considered, whereas practical systems may experience multiple simultaneous or evolving faults. In addition, the ANN performance depends on the quality and diversity of the training dataset. Variations caused by environmental conditions, sensor inaccuracies, measurement noise, and module aging may affect the classification accuracy when the model is deployed in real-world conditions. Finally, the PCA transformation, while efficient, may remove certain subtle features that could contribute to fault discrimination.

IV.9.3 Perspectives for Improvement and Future Work:

Future research can focus on enhancing the robustness and generalization capability of the proposed diagnostic framework. The integration of real experimental datasets collected from operating PV plants would provide more realistic validation of the model. Advanced Long Short-Term Memory (LSTM) networks, and hybrid AI architectures could be investigated to improve diagnostic accuracy for complex fault scenarios. Additional electrical, thermal, and environmental features may also be incorporated to enrich the fault representation. Furthermore, the implementation of the diagnostic algorithm on embedded systems, edge-computing devices, or Internet of Things (IoT) platforms would enable real-time monitoring and predictive maintenance of photovoltaic installations. Such developments would contribute to improving the reliability, efficiency, and operational lifespan of modern solar energy systems.

General Conclusion

This thesis investigated the use of artificial intelligence (AI) techniques for the diagnosis of faults in photovoltaic (PV) panels and evaluated their effectiveness in improving the performance, reliability, and maintenance of PV systems. As PV installations continue to expand worldwide, the rapid and accurate detection of faults has become essential to ensure optimal energy production and operational safety.

The study began with a review of PV system fundamentals and the principal fault mechanisms that affect panel performance, including partial shading, open-circuit faults, short-circuit faults, and degradation-related defects. A simulation model was then developed to reproduce both normal and faulty operating conditions, enabling the generation of representative data for diagnostic purposes.

Several AI-based approaches were investigated for fault detection and classification. Through data preprocessing, feature extraction, and machine learning techniques, the proposed diagnostic framework was able to identify fault conditions automatically and distinguish between different fault types. The obtained results demonstrated that AI methods can effectively analyse PV system behaviour and provide accurate fault classification, reducing the limitations associated with conventional diagnostic approaches.

The performance evaluation and comparative analysis confirmed that intelligent diagnostic models improve the speed, accuracy, and reliability of fault detection. These improvements contribute to enhanced system monitoring, reduced maintenance costs, minimised energy losses, and increased operational efficiency of PV installations.

In conclusion, the integration of artificial intelligence with photovoltaic fault diagnosis represents a promising solution for modern PV systems. The results obtained in this work demonstrate that AI-based diagnostic techniques can significantly improve the reliability and performance of photovoltaic panels while supporting predictive maintenance strategies. Future work may focus on the use of real-world datasets, advanced deep learning architectures, and real-time implementation on embedded monitoring platforms to further enhance diagnostic accuracy and practical deployment.

References:

[1]: Photosynthesis Light Sensor and Meter. Ecology, Vol. 52, pp. 125–131 Biggs, W.W., A. R. Edison, J. D. Eastin, K.W. Brown, J.W. Maranville, and M.D. Clegg. (1971).

[2]: Applied Photovoltaics /Second Edition, S.R. Wenham, M.A. Green, M.E. Watt, and R. CORKISH, first published by Earthscan in the UK and USA in 2007, Copyright © 2007, ARC Centre for Advanced Silicon Photovoltaics and Photonics

[3]: Principles of Solar Engineering, D. Yogi Goswami / CRC Press, Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742, © 2015 by Taylor & Francis Group, LLC. CRC Press is an imprint of Taylor & Francis Group, an Informa business. No claim to original U.S. government works. Version Date: 20141017. International Standard Book Number-13: 978-1-4665-6379-7 (eBook - PDF).

[4]: Solar Energy the Physics and Engineering of Photovoltaic Conversion, Technologies and Systems, Arno HM Smets, Klaus Jäger, Olindo Isabella, René ACMM van SWAAIJ, Miro Zeman

[5]: Electricity from Sunlight Photovoltaic Systems Integration and Sustainability, Second Edition, Vasilis Fthenakis, Columbia University, USA Paul A. Lynn, formerly of Imperial College London, UK, this second edition was first published in 2018, 2018 John Wiley & Sons Ltd., Edition History, John Wiley & Sons Ltd. (1e, 2010).

[6]: Handbook of Photovoltaic Science and Engineering, Second Edition, Edited by Antonio Luque, Instituto de ENERGIA Solar, Universidad Polit ECNICA de Madrid, Spain, and Steven Hegedus, Institute of Energy Conversion, University of Delaware, USA

[7]: Solar Radiation and Daylight Models (with software available from companion website) T. Muneer, Napier University, Edinburgh, with a chapter on Solar Spectral Radiation by C. Gueymard, Solar Consulting Services, Denver, Colorado, and H. KAMBEZIDIS, National Observatory of Athens, Athens, Elsevier Butterworth-Heinemann, Linacre House, Jordan Hill, Oxford OX2 8DP, 200-Wheeler Road, Burlington, MA 01803. First published in 1997. Second edition 2004, Copyright © 1997, 2004, Elsevier Ltd. All rights reserved.

[8]: Photovoltaic Systems Engineering for Students and Professionals, Designed cover image: Drawings and design by MUGDESEM TANRIOVEN. First edition published 2024 by CRC Press, 6000 Broken Sound Parkway NW, Suite 300, Boca Raton, FL 33487-2742, and by CRC Press, 4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN. Press is an imprint of Taylor & Francis Group, LLC.

[9]: Physics of Semiconductor Devices Third Edition, S. M. Sze, Department of Electronics Engineering, National Chiao Tung University, Hsinchu, Taiwan, And Kwok K. Ng, Central Laboratory, San Jose, California July 2006.

[10]: Principles of Semiconductor Devices, second edition, Sima Dimitrijev Griffith University, 2012, 2006 by Oxford University Press, Inc.

[11]: Solar energy the physics and engineering of photovoltaic conversion, technologies and systems, June 2015 Ronald J. Ch. van Zolingen Emeritus professor, Eindhoven University of Technology.

[12]: Photovoltaic Power System Modeling, Design, and Control, Weidong Xiao, University of Sydney Australia, this edition first published 2017, Set in 10/12pt WarnockPro by SPi Global, Chennai, India.

[13]: Photovoltaic System Design Procedures, Tools and Applications. Suneel Deambi, 2016 by Taylor & Francis Group, LLC CRC Press is an imprint of Taylor & Francis Group, an Informa business, No claim to original U.S. Government works, Printed on acid-free paper Version Date: 20151123, International Standard Book Number-13: 978-1-4822-5980-3 (Hardback) LC record available at http://lccn.loc.gov/2015045365.

[14]:MDPI Anti-Reflective Coating Materials: A Holistic Review from PVPerspective Natarajan Shanmugam 1 , Rishi Pugazhendhi 1, Rajvikram Madurai Elavarasan 2,* Pitchandi Kasiviswanathan 1 1 andNarottamDas3,4 , Department of Mechanical Engineering, Sri Venkateswara College of Engineering, Chennai 602117, India; aspnatraj@gmail.com (N.S.); rishi2000.p@gmail.com (R.P.); pitch@svce.ac.in (P.K.) 2 3 4 * Department of Electrical and Electronics Engineering, Sri Venkateswara College of Engineering, Chennai 602117, India School of Engineering and Technology, Central Queensland University, Melbourne, VIC 3000, Australia; n.das@cqu.edu.au Centre for Intelligent Systems, School of Engineering and Technology, Central Queensland University, Brisbane, QLD 4000, Australia Correspondence: rajvikram787@gmail.com Received: 6 April 2020; Accepted: 15 May 2020; Published: 21 May 2020.

[15]: Applied Photovoltaics Third Edition, S.R. Wenham, M.A. Green, M.E. Watt, R.P. Corkish, A.B. Sproul, First published 1994 by Centre for Photovoltaic Devices and Systems, University of New South Wales, Australia, Second edition published in the UK and USA in 2007 by Earthscan, Third edition first published 2011, by Earthscan 2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN, Simultaneously published in the USA and Canada by Earthscan

[16]: Applied Photovoltaics Second Edition, S.R. Wenham, M.A. Green, M.E. Watt, R. Corkish, First published by Earthscan in the UK and USA in 2007.

[17]: Photovoltaics System Design and Practice, Heinrich Ha¨berlin, Berne University of Applied Sciences, Switzerland Translated by Herbert Eppel HE Translations, Leicester, UK, This edition first published 2012, John Wiley & Sons, Ltd.

[18]: MPI Article Design of the Front Contact Metallization Patterns for Solar Cells Using Variable-Width Bezier Curves Kai Li 1,2,* , Yongjiang Liu 1 and Peizheng Li 1 1 School of Mechanical Engineering, Xinjiang University, Urumqi 830047, China 2 Guangdong Key Laboratory of Precision Equipment and Manufacturing Technology, South China University of Technology, Guangzhou 510641, China * Correspondence: scutlk@163.com.

[19]: Principles of Solar Cells, LEDs and Diodes the role of the PN junction, ADRIAN KITAI Departments of Engineering Physics and Materials Science and Engineering, McMaster University, Hamilton, Ontario, Canada, This edition first published 2011, John Wiley & Sons, Ltd

[20]: Photovoltaic Module Reliability, John H. Wohlgemuth Virginia USA, this edition first published 2020 John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK For

details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

[21]: Solar Energy: Advancements and Challenges / Gaurav Saini, Korhan Cengiz, Sesha Srinivasan, Sanjeevikumar Padmanaban and Krishna Kumar. 2023 River Publishers. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, mechanical, photocopying, recording or otherwise, without prior written permission of the publishers.

[22]: Silicon Solar Cell Metallization and Module Technology, Edited by Thorsten Dullweber and Loic Tous. The Institution of Engineering and Technology Published by The Institution of Engineering and Technology, London, United Kingdom. The Institution of Engineering and Technology is registered as a Charity in England & Wales (no. 211014) and Scotland (no. SC038698). the Institution of Engineering and Technology 2022 First published 2021.

[23] J. A. Duffie and W. A. Beckman, *Solar Engineering of Thermal Processes*, 4th ed. Hoboken, NJ, USA: John Wiley & Sons, 2013.

[24] G. R. Walker, "Evaluating MPPT converter topologies using a MATLAB PV model," *Journal of Electrical and Electronics Engineering, Australia*, vol. 21, no. 1, pp. 49–56, 2001.

[25] J. A. Duffie and W. A. Beckman, *Solar Engineering of Thermal Processes*. New York, NY, USA: John Wiley & Sons, 2013.

[26] R. A. Messenger and J. Abtahi, *Photovoltaic Systems Engineering*, 4th ed. Boca Raton, FL, USA: CRC Press, 2017.

[27] M. H. Rashid, Power Electronics: Circuits, Devices, and Applications, 4th ed. Upper Saddle River, NJ, USA: Pearson Education, 2017.

[28] M. G. Villalva, J. R. Gazoli, and E. Ruppert Filho, "Comprehensive approach to modeling and simulation of photovoltaic arrays," IEEE Transactions on Power Electronics, vol. 24, no. 5, pp. 1198–1208, 2009.

[29] H. Patel and V. Agarwal, "MATLAB-based modeling to study the effects of partial shading on PV array characteristics," IEEE Transactions on Energy Conversion, vol. 23, no. 1, pp. 302–310, 2008.

[30] M. Seyed mahmoudian, B. Horan, R. Rahmani, A. M. T. Oo, A. Stojcevski, and A. Mekhilef, "State of the art artificial intelligence-based MPPT techniques for mitigating partial shading effects on PV systems—A review," Renewable and Sustainable Energy Reviews, vol. 64, pp. 435–455, 2016.

[31] K. Ishaque and Z. Salam, "A review of maximum power point tracking techniques of PV system for uniform insolation and partial shading condition," Renewable and Sustainable Energy Reviews, vol. 19, pp. 475–488, 2013.

[32] A. Chouder and S. Silvestre, "Analysis model of mismatch power losses in PV systems," Journal of Solar Energy Engineering, vol. 132, no. 2, pp. 024504-1–024504-6, 2010.

[33] E. Garoudja, K. Chouder, S. Kara, A. Silvestre, and E. Chouder, “An enhanced machine learning based approach for photovoltaic fault detection and diagnosis,” Energy Conversion and Management, vol. 151, pp. 634–648, 2017.

[34] M. Dhimish, V. Holmes, B. Mehrdadi, and M. Dales, “Simultaneous fault detection algorithm for grid-connected photovoltaic plants,” IET Renewable Power Generation, vol. 11, no. 12, pp. 1565–1575, 2017.

[35] W. Chine, A. Mellit, A. M. Pavan, and V. Lughi, “Fault diagnosis in photovoltaic arrays using electrical parameters and artificial neural networks,” Renewable Energy, vol. 90, pp. 501–512, 2016.

[36] A. Mellit, G. Tina, and S. A. Kalogirou, “Fault detection and diagnosis methods for photovoltaic systems: A review,” Renewable and Sustainable Energy Reviews, vol. 91, pp. 1–17, 2018.

[37] Gazoli J.R., and Filho E.R., “Comprehensive Approach to Modeling and Simulation of Photovoltaic Arrays,” *IEEE Transactions on Power Electronics*, vol. 24, no. 5, pp. 1198–1208, 2009.

[38] Photovoltaic Systems Engineering, Roger A. Messenger and Jerry Ventre, CRC Press, 2017.

[39] Photovoltaic Engineering Handbook, Wiley, 2011.

[40] International Energy Agency, *Review of Failures of Photovoltaic Modules*, IEA-PVPS Report.

[41] Photovoltaics literature on PV fault diagnosis and performance assessment.

[42] *Pattern Recognition and Machine Learning*. Springer, (2006).

[43] Ian T. Jolliffe (2016). *Principal Component Analysis* (2nd ed.). Springer.

[44] Jiawei Han, Kamber M., & Pei J. (2022). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.

[45] Charu C. Aggarwal (2015). *Data Mining: The Textbook*. Springer.

[46] Stephane Mallat (2009). *A Wavelet Tour of Signal Processing*. Academic Press.

[47] Abdelhamid Mellit & Spyros A. Kalogirou (2014). Artificial intelligence techniques for photovoltaic applications: A review. *Progress in Energy and Combustion Science*, 41, 92–117.

[48] IEEE publications on photovoltaic monitoring, feature extraction, and fault diagnosis.

[49] Rodrigues et al. (2022). Machine-learning-based fault detection and diagnosis in photovoltaic systems: Recent advances and challenges. *Renewable and Sustainable Energy Reviews*.

[50] International Electrotechnical Commission (2021). IEC 61724-1: Photovoltaic System Performance Monitoring – Guidelines for Measurement, Data Exchange and Analysis.

[51] Madeti Srinivasa Rao & Singh S. N. (2017). Monitoring Systems for Photovoltaic Plants: A Review. *Renewable and Sustainable Energy Reviews*, 67, 1180–1207

[52] Artificiel Intelligence Islam, M. R., et al. (2023). *Artificial Intelligence in Photovoltaic Fault Identification and Diagnosis: A Systematic Review*. Energies, 16(21), 7417.

[53] Machine Learning Yuan, Z., Xiong, G., & Fu, X. (2022). *Artificial Neural Network for Fault Diagnosis of Solar Photovoltaic Systems: A Survey*. Energies, 15(22), 8693.

[54] Deep Learning LeCun, Y., Bengio, Y., & Hinton, G. (2015). *Deep Learning*. Nature, 521(7553), 436–444.

[55] Ian Jolliffe (2016). *Principal Component Analysis* (2nd ed.). Springer.

[56] Khan, F., Kim, J., & co-authors (2021). Deep learning-based fault detection and diagnosis techniques for photovoltaic systems: A review. *Renewable Energy*.

[57] Rodrigues, E. M. G., Melício, R., Mendes, V. M. F., & Catalão, J. P. S. (2022). Machine learning-based fault detection and diagnosis in photovoltaic systems: Recent advances and challenges. *Renewable and Sustainable Energy Reviews*.

[58] Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). *Edge Computing: Vision and Challenges*. IEEE Internet of Things Journal, 3(5), 637–646.

[59] bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

[60] Cortes, C., & Vladimir Vapnik (1995). Support-Vector Networks. *Machine Learning*, 20(3), 273–297.

[61] Goodfellow, I., Yoshua Bengio, & Aaron Courville (2016). *Deep Learning*. MIT Press.

[62]). *Pattern Recognition and Machine Learning*, Christopher M. Bishop (2006. Springer.

[63] Nitish Srivastava, Hinton, G., Krizhevsky, A., Sutskever, I., & Sala khutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15, 1929–1958.

[64] IEEE publications on deep learning applications in photovoltaic fault diagnosis and renewable energy systems.

[65] Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521, 436–444.