Intelligent Fault Management for Photovoltaic System using Cross-Correlation and Machine Learning

C. U. Nwamuo[1], O. Ojike[2,*], S. O. Enibe[3]

[1]Renewable and New Energy System, African Center of Excellence for Sustainable Power and Energy Development, University of Nigeria Nsukka, ACESPED, Enugu, Nigeria.
[2]Bio Resources Engineering Department, University of Nigeria Nsukka, Enugu, Nigeria.
[3]Mechanical Engineering Department, University of Nigeria Nsukka, Enugu, Nigeria.

ABSTRACT: This paper presents a fault management model for the protection of off-grid PV system. Data was collected from an off-grid 4.2 kW PV system installed at Edugen technologies Ltd, located at Nsukka, Enugu State, Nigeria. An adaptive feature extractor using a cross-correlation approach was used to identify and extract intricate and high dimensional features of PV characteristics. This data was then used to train a neural network. For the training of the network, Bayesian optimization (Bayes) and adaptive movement estimation (Adam), were used individually and as hybrid. The models generated were evaluated experimentally, considering Mean Square Error (MSE), Coefficient of determination (R^2), and Root Mean Square Error (RMSE). Comparative analysis of the three training algorithms showed that the hybrid optimization algorithm recorded better R^2 of 0.9824, MSE of 9.5673e^{-9}, and RMSE of 9.781e^{-5} when compared to Bayes and Adam. In conclusion, the hybrid optimization-based PV fault detection model was the most suitable for PV protection and is therefore recommended for the protection of off-grid standalone PV systems to ensure power sustainability.

KEYWORDS: Photovoltaic, Bayesian, Adam, hybrid optimization, fault, cross correlation

I. INTRODUCTION

Globally, renewable energy sources are playing critical roles in the electricity generation sector. Notably, the Photovoltaic (PV) constitutes one of the most promising for green energy technologies due to several advantages such as global availability, modularity, ease of implementation, low maintenance cost, and environmental friendliness (El-Banby et al., 2023). Since 2000, the global cumulative installed PV capacity has continued to increase on a yearly basis, according to Al-Katheri et al. (2022). For instance, in 2020, global PV installed was 773.2 GW, which was 138 GW more than the previous year (Global Cumulative Installed Solar, 2019). Nevertheless, Garoudja et al. (2017) and Tina et al. (2016) revealed that PV systems are prone to several types of faults, which impact their reliability.

According to the United Kingdom (UK) energy audit by Firth et al. (2010), faults in PV systems result in 18.9% loss annually, thus necessitating the need for a Fault Detection System (FDS). Faults in PV originate from various parts of the system, which range from the panel fault, module aging, shadow fault, and dirty modules and can be generally classified into physical, electrical, and environmental faults as posited by Garoudia et al. (2017). The physical faults occur mainly on the panel due to damage on the cell, diode bypass, cracks on the panel, or degradation due to age factors. Environmental faults are external issues such as shading faults, which occur because of cloud movement, tree shadows, and leaves falling on the panel and blocking input of irradiance. Electrical faults are another type of fault Yuan et al. (2022) classified into array faults and power conditioning unit faults. While the latter are divided into Direct Current (DC) and Alternative Current (AC) faults, the former comprise line-to-line, open circuit, and line-to-earth faults. Batteries, maximum power point trackers, and switch converters are examples of DC devices that can experience DC faults, whereas grid irregularities and islanding operations are examples of AC faults.

Generally, the methods of fault detection are classified by Tina et al. (2016) into electrical, thermal, and visual methods. The electrical method focuses on current-voltage (IV) characteristics curve analysis, Power Loss Analysis (PLA), Voltage and Current Measurement (VCM), Artificial Intelligence Techniques (AIT), and Signal Processing (SP) according to Mellit et al. (2018) and El-Rashidy (2022), while the thermal and visual method include techniques such as florescence imaging, ultraviolet technique (Kontges et al., 2014), physical inspection, and thermograph techniques (El-Banby et al., 2023).

Today, Machine Learning (ML) has continued to dominate studies for fault detection. ML is AIT which can learn from data to solve prediction or classification problems, according to Ebere et al. (2025). In the context of faults in PV systems, Utama et al. (2023) used the three Explainable Artificial Intelligence (XAI) approaches: Shapley Additive Explanations (SHAP), Anchors, and Diverse Counterfactual Explanations (DiCE). The model's comparative analysis showed that SHAP with MLNN had the highest classification accuracy, at 99.11%. In a different study Zhang et al. (2022), the fault characteristics of mismatched PV strings were determined using the I-V curve. The technique builds a detection line from the I-V data using the line parallel movement and the positional relationships to identify inflection steps on the curve. Out of 200 mismatch

*Corresponding author: chinacunch@gmail.com
doi: *http://doi.org/10.63746/njtd.v22i4.3085*

samples applied to test the model, 100% recall was reported and 96.6% precision was also recorded. Stack ensemble learning and ANN were combined in Mellit et al. (2023a) for the classification of fault data, considering attributes such as current, voltage, irradiance, and temperature, and recorded 96.50% classification accuracy when tested with a shade fault. The model was deployed in a microcontroller and integrated with the Blynk Internet of Things (IoT) dashboard application as an embedded system for remote monitoring of PV faults. Wang et al. (2023) combined sparse autoencoder and Bayesian inference for the calibration of PV fault and recorded Root Mean Square Error (RMSE) of 0.0014. While these papers all contributed successfully to PV fault detection, Rouani et al. (2021) argued that PV data are highly uncertain and require a model that is proactive to model this behaviour. Vertices Principal Component Analysis (VPCA) was proposed and trained considering shading fault data. However, the 35.7% false detection rate recorded leaves room for improvement.

In Sairam et al. (2022), an irradiance-based three-diode panel fault classifier was developed using an extreme gradient boosting algorithm. The results after evaluation using RMSE reported 0.00799. Lakshmi et al. (2023) also proposed a combination of a light gradient boosting algorithm and a decision tree. This hybrid model was trained with input PV data containing current, voltage, and temperature for the prediction of fault. After evaluation, MSE reported 8.84, while coefficient of determination R^2recorded 0.9939. 90% classification success was reported in Li et al. (2019), who used ANN to detect different types of faults, which include electrical and shadow faults, while in Et-Taleby et al. (2023), deep learning algorithms Convolutional Neural Network (CNN), Region-based CNN (R-CNN), Ensemble algorithm (ANN, Deep-CNN, decision tree), and machine learning (Probabilistic Neural Network (PNN), Principal Component Analysis (PCA), Support Vector Machines (SVM), and K-Nearest Neighbours (k-NN)) were analysed for the detection of PV fault. Results after evaluation with solar panel failure data reported an accuracy of 99.49% for K-NN, SVM, and CNN; PNN reported 100%, while DT and ANN reported 99.65%.

According to several studies, deep learning models have outperformed traditional machine learning in various applications due to their ability to learn high dimensional data better (Hong and Pula, 2022; Chidi et al., 2024; Kekong et al., 2019; Dai and Salim, 2020; Deng et al., 2014). In the context of PV faults, Liu et al. (2021) combined stacked autoencoder and clustering approaches for the detection and diagnosis of PV faults. Chen et al. (2019) trained a Deep Residual Network (DRN) using I–V curve characteristics and ambient temperature data for intelligent fault diagnosis. CNN and residual gated recurrent units were applied by Gao and Wai (2020) to monitor deviations in the I–V curve and classify faults in PV systems. Similarly, Appiah et al. (2019) combined Long Short-Term Memory (LSTM) and SoftMax regression classifiers for the detection and diagnosis of faults using I–V characteristic data. In Liu et al., (2023), a comparative analysis of Infrared Thermography (IRT) considering integrated Deep Learning (DL) and ML. The findings showed the DL approach recorded better classification success than the traditional ML. The reason according to Deng (2014) was due to DL ability to learn from high-dimensional data. Meanwhile Sarker (2021) revealed that DL suffers several challenges which include high data demands, high resource computation requirements, being prone to over-fitting with small datasets, and struggling to adapt to changing real-world problems. In addition, DL requires a Graphical Process Unit (GPU), which in most cases is expensive to subscribe to.

Multi-Layer Perceptron (MLP) provides an alternative approach to address the challenges faced by deep learning algorithms (Al-Katheri et al., 2022). MLP is a type of feed-forward neural network with multiple interconnected layers that learn to solve complex problems. Several studies have shown that MLPs are robust to noisy data, offer high accuracy, and are simple yet efficient, with fewer resource requirements and a reduced risk of overfitting—particularly with limited datasets. They are also flexible, as they can be easily retrained (Al-Katheri et al., 2022; Chen et al., 2020; Sarker, 2021; Habor et al., 2021). Mekki et al. (2016) applied MLP for the estimation of current and voltage variation due to solar irradiance and cell temperature. Chine et al. (2016) considered eight different types of faults using I-V characteristics data with inputs of open circuit voltage and current to train MLP. Rao et al. (2019) applied SoftMax function in the output layer of MLP or the classification of PV faults considering I-V characteristics. Ul-Haq et al. (2020) also combined scaled conjugated gradient and MLP to train with PV fault data collected from poly-crystalline PV technologies, while Khelil et al. (2021) applied back-propagation algorithm and MLP for the detection of short circuit faults in PV systems. Gap resides in the need to consider other types of faults, such as environmental and physical fault. In addition, the need for an adaptive model capable of learning dynamic and heterogeneous behaviour of PV fault conditions has already been identified. Therefore, modelling a PV fault detection system that considers these problems presents the need for this research. The major contributions of this paper are as follow.

i. To propose a cross-correlated feature extraction model which identifies the intricate features of PV fault data
ii. To propose an adaptive optimization algorithm for improved model convergence, optimal hyper-parameter selection and generalization
iii. To compare the performance of neural network on three different training algorithms and validate the best with real fault data.

II. THEORITICAL ANALYSIS

This section presents theory of Photovoltaic (PV) technology and types of PV systems.

a. Photovoltaic Technology

Photovoltaic (PV) technology, commonly known as solar technology, involves transforming energy from the sun into electricity using semiconductor materials (El-Banby et al., 2023). It is the mainstay of photovoltaic technology and constitutes mainly semiconductor materials like silicon. Semiconductor materials, including silicon, are used to make solar cells. When light strikes the cell, the energy of photons of light is transferred to the charge carriers. The electric field

across the junction separates the positive charge carriers created by photons (holes) from their negative counterparts, electrons. In this manner, once the circuit is closed on an external load electrical current is extracted (Weidong, 2017).

b. Types of PV Systems

There are three major connections of PV systems which are Stand-alone (off-grid), Grid connected and Hybrid system (Heinrich, 2016).

i. Stand-alone (off-grid)

A Stand-Alone Photovoltaic (PV) System, also known as an off-grid PV system, is an independent solar energy system that operates autonomously and is not connected to the electrical grid. This type of installation is basically used in remote areas where grid access is impractical or unavailable (Heinrich, 2016). Stand-alone PV systems are designed to generate, store, and deliver electricity to serve the energy demands of a specific usage. These systems comprise of solar panels, charge controllers, battery bank, inverters, DC loads and AC loads.

ii. Grid connected

An electrical grid connection enables the smooth integration of solar-generated electricity with the utility grid (Kachhwaha et al., 2022). This type of system is known as a grid-connected photovoltaic (PV) system. These systems are majorly employed across all home categories to harness solar energy, contributing to the overall electricity supply. This system also supplies auxiliary power during any short period of grid failure (El-Banby et al., 2023). For frequency control, voltage and waveform in order to meet the requirements for feeding power into grid, this system needs additional equipment. The basic components that are seen in a grid connected PV system include: the PV arrays, inverters, mounting structure grid connection and inverter standards, monitoring and control system, net metering, AC loads. Grid-connected PV systems offer benefits such as reduced electricity bills, environmental sustainability, and the potential for financial incentives.

iii. Hybrid systems

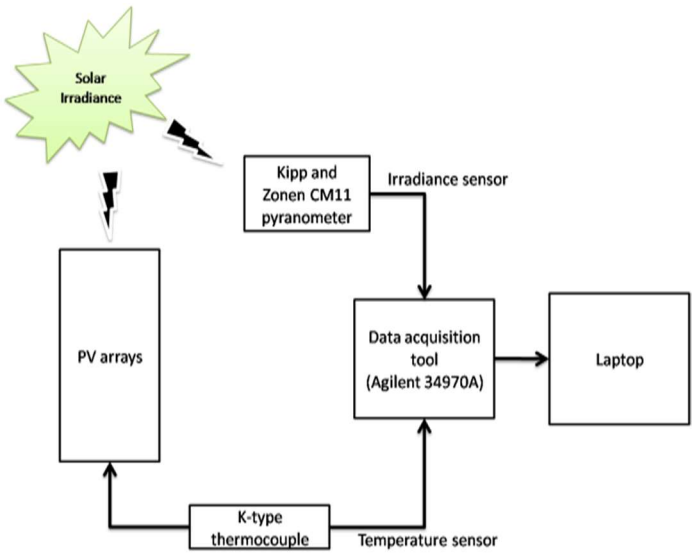
Photovoltaic modules combined with an additional power source, like a wind or diesel generator, make up hybrid systems. For hybrid PV systems, more sophisticated controls are typically needed to optimize the combined output of the two generators compared to standalone ones (Kachhwaha et al., 2022). The diesel engine in PV/diesel systems, for example, needs to start when the battery gets to a specific discharge level and equally stop when the battery gets to an adequate state of charge.

III. MATERIALS AND METHOD

The methodology of this work began with data collection of PV fault records, considering several fault types such as earth fault, shadow fault, open circuit fault, and short circuit faults from a PV installation. The collected data were processed to address duplicated values and missing values before the application of a proposed feature extraction technique with the capability to identify and extract dynamic and heterogeneous PV fault attributes, which were then applied to train a MLP neural network. During the training, a proposed adaptive optimization algorithm was applied for optimal hyper-parameter selection and model convergence, and then dropout techniques were also applied to ensure model generalization of the fault detection network. The generated model was validated through comparison with other training algorithms such as Adam and Bayesian optimization respectively.

A Data Collection

The data used for this work was collected from an off grid (12panel * 350watt) 4.2kW PV power system installed at Edugen technologies Ltd, located at Nsukka, Enugu State, Nigeria. The main parameters considered for data collection are temperature, irradiance, current and voltage. The duration for the data collection is 13hr/daily (6am to 7pm), for 11 months (January to November 2023). The main instruments used for the data collection are clamp meter, digital mustimeter, Agilent 34970A data logger and Hp Laptop. Figure 1 shows the experimental setup and PV system layout. Figure 1(a) presents the experimental setup used for the data collection while the Figure 1(b) showed the testbed. In the setup in Figure 1(a), a Kipp and Zonen CM11 pyrometer was used to measure solar irradiance on the PV panels, while K-type thermocouple was used for the measurement of the PV temperature. Both sensors feed the collected data to the Agilent 34970A data logger tool which records the data and relay to the laptop for analysis. Fault condition such as open circuit fault, short circuit fault, earth fault and shadow faults were induced manually through to alteration of PV array connections as shown in Figure 1b. For instance, the open circuit fault was induced through the disconnection of PV circuits to create the open circuit voltage and zero current which was measured with the data logger tool. Two of the PV arrays were connected to the ground to induce earth fault, while to induce short circuit fault, one of the panels were disconnected and used for this test, by connecting the positive and negative terminal together, while the data logger was used to measure the I-V characteristics. To simulate shadow fault, part of the panel was covered, while the data logger measures the I-V characteristics as it fluctuates.



(a)

(b)

Figure 1 Experimental Setup and PV System Layout (a) The experimental setup used for data measurement (b) Layout of a PV power system showing the separation and organization of cables from the PV module, and the integrated setup of the inverter and battery bank for power conversion, storage, and distribution.

B. Modelling of the PV System

A PV system is made of solar cells which consist of semiconductor materials such as silicon. During the operation of the system photons from the sun are absorbed by the solar cell, creating electron-hole pairs, and producing a potential difference across the cell, which, when connected to an external circuit, causes current to flow. The behaviour of the solar cell in varying temperatures and amounts of sunlight is determined by its electrical properties. The electrical behaviour of a PV module is typically captured using an equivalent circuit model. This model represents the various components of the PV module and their interrelationships. Figure 2 illustrates the typical components of the equivalent circuit, which include a Series Resistance (R_s), a shunt resistance (R_{sh}), a diode that represents the pn junction (I_d), and a current source or photocurrent (I_{ph}). The current source is in parallel with the diode. Also, $R_{sh}$and R_s are connected in parallel and series respectively with the current source.

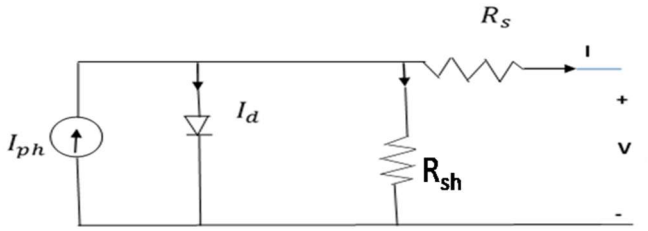


Figure 2 Equivalent circuit of PV cell (Livera et al., 2019)

By applying Kirchhoff's current law in Figure 2, the total PV current, I, is given in Equation 1.

$$I = I_{ph} - I_d - I_{sh} \quad (1)$$

Where the photocurrent (I_{ph}) is measured when the photovoltaic cell is exposed to incident sunlight. At a given temperature, the photocurrent exhibits linear variation and is dependent upon the solar irradiance. The non-linear response on the PV cell is caused by I_d, an anti-parallel diode current.I_{sh} is the representation of the current flowing through the shunt resistor. Equation 2 yields the PV current when the expressions for $I_{sh}$and I_d in Equation 1 are changed.

$$I = I_{ph} - I_0\left[\left\{e^{\frac{q(V+IR_S)}{AKT}}\right\} - 1\right] - \frac{(V+IR_S)}{R_{sh}} \quad (2)$$

Where, I_o = Module saturation current, V= solar cell voltage, A= ideality factor, K= Boltzmann constant, q= electron charge, T= temperature of the PV cell. The estimation of the thermal voltage of a PV module under partial shading conditions has been expressed by Equation 3;

$$V_{te} = \frac{(2V_{mpp}-V_{oc})(I_{sc}-I_{mpp})}{I_{mpp}-(I_{sc}-I_{mpp})\mathrm{In}(\frac{I_{sc}-I_{mpp}}{I_{sc}})} \quad (3)$$

Where, V_{mpp}=voltage at maximum power point, I_{mpp}= current at the maximum power point, V_{oc}= open circuit voltage, and I_{sc}=short circuit current. Equation 3 when applied under partial shading condition is presented as Equation 4;

$$V_{te} = \frac{N_s \mathrm{AkT}}{\mathrm{q}} \quad (4)$$

Where, N_s is the number of solar cells connected in series, A is the diode ideality factor, k is the Boltzmann constant ($1.381 * 10^{-23} J/K$) and T is temperature and q = $1.6021 * 10^{-19} Coloumb$, is the charge. The PV's photocurrent is influenced by both the temperature of the cell and the amount of irradiation. Equation 5 provides the photocurrent expression.

$$I_{ph} = [I_{sc} + K_i(T - 298)]x\frac{I_r}{I_{ro}} \quad (5)$$

Where, K_i= short-circuit current temperature coefficient of cell (0.0024); I_{r0}= solar irradiation constant (1000 W/m^2 at Standard Reference Condition). The module saturation current (I_0) varies with the cell temperature, which is given by Equation 6.

$$I_o = I_{rs}[\frac{T}{T_r}]^3 \exp\left[\frac{qxE_g}{nk}\left(\frac{1}{T} - \frac{1}{T_r}\right)\right] \quad (6)$$

Where, T = PV cell temperature, T_r = Standard Reference Temperature (T_r= 298 K or 25˚C), I_{rs}= modules reverse saturation current, E_g = Band gap energy.

$$E_g = E_{g0} - \alpha\frac{T^2}{T+\beta} = 1.6 - 7.02\,x10^{-4}[\frac{T^2}{T-1108}] \quad (7)$$

The modules reverse saturation current (I_{rs}) can be expressed using Equation 8.

$$I_{rs} = \frac{I_{sc}}{e^{\frac{q(V+IR_S)}{nKT}-1}} \quad (8)$$

PV cells are large, grouped units known as PV modules. The diodes in these modules are connected in series or parallel to form PV arrays, which are used in PV generation systems to produce electricity. Figure 3 depicts the equivalent circuit for a photovoltaic array.

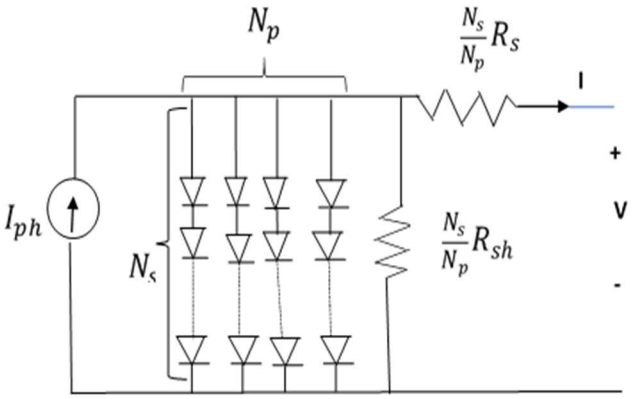


Figure 3 Equivalent Circuit of Solar Array (Livera et al., 2019)

The solar cells should be arranged in series and shunt combinations to maximize the capacity of solar PV systems to generate electricity. The relationship between the output current and voltage is given as follows Equivalent Circuit of Solar Array N_s for the cells connected in series.

$$I = N_p x I_{ph} - N_p x I_0 x \left[\exp\left(\frac{{}^{V}/_{R_s}+Ix^{R_s}/_{N_p}}{nxV_{te}}\right)\right] - I_{sh} \quad (9)$$

$$I_{sh} = \frac{Vx^{N_p}/_{N_s+IxR_s}}{R_p} \quad (10)$$

The power-voltage (P-V) curve, which depicts the relationship between power (P) and voltage (V), can be used to describe the power output of a solar cell. The solar cell's power output is given by:

$$P = V * I \quad (11)$$

The maximum power point (MPP) of the solar cell, where the power output is maximized, can be calculated as:

$$P_{max} = I_{SC} * V_{OC} * FF \quad (12)$$

Where: I_{sc} is the short circuit current (A), V_{OC} is the open − circuit voltage (Volts), *FF*= the fill factor (dimensionless). The Fill Factor (FF) is a measure of the solar cell's efficiency in converting sunlight into electricity. It is represented as the relationship between the maximum power output and the sum of the short- and open-circuit currents. It is given below:

$$FF = \frac{I_{mpp}V_{mpp}}{I_{SC}*V_{OC}} \quad (13)$$

Sunlight irradiance is the amount of electromagnetic radiation power per unit area that is received from the Sun. This parameter holds great importance in multiple domains, including atmospheric science, climate modelling, and solar energy production. An individual's location on Earth, the time of day, the season, the state of the atmosphere, and other factors all affect how much solar radiation they receive. There are roughly 1361 watts per square meter (W/m^2) of solar radiation that reaches Earth's outer atmosphere; this is also referred to as the solar constant. However, this value varies slightly due to factors such as solar cycles and orbital variations. To calculate solar irradiance at a specific location on Earth's surface, several factors need to be considered, including the solar zenith angle (θ), atmospheric attenuation, and cloud cover. The mathematical representation of solar irradiance at a given location and time is given as:

$$I_r = I_{ro}.\cos(\theta)\, e^{-\tau/\mathrm{co}\ (\theta)} \quad (14)$$

Where I_r = solar irradiance on a horizontal surface at the Earth's surface (W/m²), Ir_0=solar constant (approximately 1361 W/m²), θ=solar zenith angle (the angle between the direction of the Sun and the vertical direction), τ = the optical depth of the atmosphere. Equation 14 considers the attenuation of solar radiation as it passes through the Earth's atmosphere. The term $cos(\theta)$ represents the geometrical reduction in solar radiation due to the angle of incidence, and $e^{-\tau/cos}$ accounts for atmospheric attenuation. The solar zenith angle (θ) varies throughout the day and depends on the latitude and time of year. At solar noon (when the Sun is directly overhead), the solar zenith angle is 0°, and as the Sun moves away from zenith, the solar zenith angle increases. The optical depth (τ) accounts for the absorption and scattering of solar radiation by atmospheric constituents such as water vapor, aerosols, and gases like ozone and carbon dioxide. It varies with atmospheric conditions, including air mass, altitude, and the presence of clouds and pollutants.

C. Data Preparation and Proposed Feature Extractor Model

The data processing steps on the collected data of PV fault was first visualized in excel spreadsheet and then automatically fixing feature duplicate issues, then after saving and exported into MATLAB environment, Multiple Imputation Approach (MIA) (Lee and Yun, 2024; Lee and Jeon, 2009; Lee, 2020) was applied to replace missing values. Traditional single imputation which used mean, mode or median of the existing values to estimate the missing values suffers issues of bias (Lee, 2020) and may affect the integrity of the completed data model generated. To ensure a more reliable and complete data quality, a MIA was adopted due to its ability to estimate multiple single imputation values and amalgamate into a unified value to replace the missing samples in the dataset. PV data are very dynamic (Wang et al., 2016; Mouaad et al., 2024) due to the time variant characteristics of current and voltage based on several factors such as temperature, irradiance, and diode interactions (Ibrahim et al., 2020). In addition, PV fault data are heterogeneous, highly dimensional and present a clustering similarity problem, which makes it very difficult for learning when directly fed to any ML algorithm. Cross Correlation Extractor (CCE) (Stähler et al., 2011; Zhang et al., 2012) is a feature extraction technique which can measure heterogeneous features which are similar, time varying data and hence make it suitable as the feature extraction of choice for this work. Traditional CCE (Bose et al., 2016; Rafique and Rafique, 2016;

Adrian-Martinez and Ardid, 2015; Camachi et al., 2018) despite their success may not be able to fully capture the complex dependencies in PV fault system under varying conditions, thus necessitating the need for an improved adaptive CCE for optimal extraction of PV signals. The proposed adaptive CCE is made of several components in Figure 4.

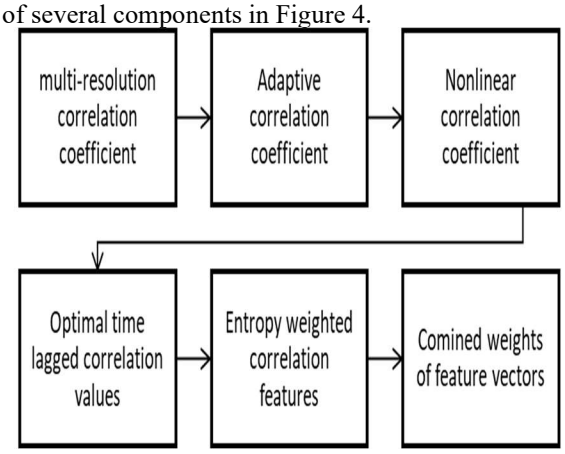


Figure 4: Block diagram of the proposed adaptive CCE

Figure 4 presents the proposed adaptive CEE components and how they interact to extract PV fault data. The Multi Resolution Correlation Coefficient (MRCC) is applied for the decomposition V(t) and I(t) representing voltage and current signals of the PV into multiple cross correlation at each scale to capture both global and local patterns using Discrete Wavelet Transform (Chen et al., 2017; Begum et al., 2024). Let $V_i(t)$ and $I_i(t)$ represent voltage and current with their i-th coefficient which was transformed with DWT as in Equation 15 and 16 respectively.

$$V(t) \underset{\rightarrow}{\overset{DWT}{}} \{V_i(t), V_2(t), \dots\dots\dots\dots V_n(t)\} \quad (15)$$

$$I(t) \underset{\rightarrow}{\overset{DWT}{}} \{I_i(t), I_2(t), \dots\dots\dots\dots I_n(t)\} \quad (16)$$

Computing the cross correlation at each scale i is computed as:

$$C_{MRCC\,(i)} = \frac{t^{(V_i(t)-\bar{V}_i)(I_i(t)-\bar{I}_i)}}{\sqrt{t^{(V_i(t)-\bar{V}_i)^2}\, t^{(I_i(t)-\bar{I}_i)^2}}} \quad (17)$$

Where $\bar{I}_i$ and $\bar{V}_i$ are the mean value of the wavelet coefficient at scale i, which gives the set of MRCC for i as the wavelet decomposition. To make the correlation adaptive, external factors such as temperature (T) and irradiance (G) are considered with adaptive weights $\alpha(t) and\, \beta(t)$ which continued to vary based on environmental factors and are presented for T and G in Equation 18 and19, while the adaptive cross correlation becomes Equation 20.

$$\alpha(t) = f1(T(t), G(t)) \quad (18)$$

$$\beta(t) = f2(T(t), G(t)) \quad (19)$$

$$C_{ACC}(t) = \frac{t^{\alpha(t)(V(t)-\bar{V})(I(t)-\bar{I})}}{\sqrt{t^{(\alpha(t)V(t)-\bar{V}_i)^2}\, t^{\beta(t)(I(t)-\bar{I})^2}}} \quad (20)$$

Equation 20 shows the variability of temperature and irradiance on the $C_{MRCC\,(i)}$. This signal is transformed into nonlinear radial kernel function K(x,y), with the nonlinear function for current denoted as φ(I(t)) and ɖ(V(T)) for voltage. Where:

$$ɖ(V(t)) = \exp -\gamma\left|\left|V(t) - V(t')\right|\right|^2 \quad (21)$$

$$\varphi(I(t)) = \exp -\gamma\left|\left|I(t) - I(t')\right|\right|^2 \quad (22)$$

Applying Equations 21 and 22 to Equation 17 produced the non-linear cross correlation model in Equation 23.

$$C_{NCC}(t) = \frac{t^{ɖ(V(t))-ɖ(\overline{V(t)})\,(I(t)-\bar{I})}}{\sqrt{t^{(ɖ(V(t))-ɖ\bar{V})^2}\, t^{\varphi(I(t))-\overline{\varphi(I)}^2}}} \quad (23)$$

Equation 23 presents the features of the PV in nonlinear form, while to extract the time varying signal as it updated and changed continuously, the time lag component τ is introduced to Equation 23 to produce Equation 24.

$$C_{TLCC}(t) = \frac{t^{(V(t+\tau))-(\bar{V})\,(I(t)-\bar{I})}}{\sqrt{t^{((V(t+\tau))-\bar{V})^2}\, t^{(I(t))-(\bar{I})^2}}} \quad (24)$$

Where the optimal time lag is defined as $\tau^* = \underset{\tau}{argmax}\, C_{TLCC(\tau)}$.

Equation 24 computes the cross correlation across different time lags to determine the time delay dependences of the current and voltage signals. To highlight the signal irregularities which are usually caused by fault conditions, the cross entropy or the defined as $H_v(t)$ for voltage and $H_I(t)$ for curent were applied to determine the weight correlated irregular features as Equation 25.

$$C_{ECC(T)} = \frac{t^{(H_v(t)(V((t))-(\bar{V})\;\; H_I(t)(I(t)-\bar{I})}}{\sqrt{t^{(H_v(t)(V(t))-\bar{V})^2}\, t^{H_I(t)(I(t))-(\bar{I})^2}}} \quad (25)$$

Combining the weights w_n of the different correlated features extracted as in Equation 26 presents the aggregate feature vector for the classification of PV fault.

$$C_{final}(t) = w_1 C_{MRCC}(t) + w_2 C_{ACC(T)} + w_3 C_{ACC}(t) + w_4 C_{TLCC(t)} + w_5 C_{ECC(T)} \quad (26)$$

Algorithm: Proposed adaptive feature extractor

1. Start
2. Identify PV I-V characteristics data at t
3. Decomposition of I-V signal over time (t) and compute the $C_{MRCC\,(i)}$
4. Apply $\alpha(t) and \beta(t)$ for adaptation of temperature and irradiance
5. Compute the adaptive cross correlation $C_{ACC}(t)$
6. Apply kernel function for signal transformation into nonlinear space
7. Compute the cross-correlation coefficient $C_{NCC}(t)$
8. Introduce the time lag function τ and check for $\tau^* = \underset{\tau}{argmax}\, C_{TLCC(\tau)}$
9. Compute $C_{TLCC}(t)$ for τ^*
10. Get the entropy $H_v(t)$ and $H_I(t)$; then determine the cross weights $C_{ECC(T)}$
11. Combine all correlated feature weights
12. Return as $C_{final}(t)$
13. End

D. The MLP Modelling and Proposed Training Algorithm MLP was used for this paper and trained using the feature extracted data to generate model for the detection of fault. The MLP is made of neurons with input vectors defined as $X = [x_1, x_2, \ldots\ldots\ldots\ldots x_n]$, the sum of the input weights are limited with an activation function $f(.)$ which in this work we applied sigmoid function, while the network after training produced target output $Y_T = [y_{T1}, y_{T2}, \ldots\ldots\ldots\ldots y_{Tn}]$. Mathematically, Bishop (1995) define neural network layer as $(n - h_1, \ldots h_l \ldots\ldots\ldots\ldots h_{L^l})$; where n is the inputs, h_l is the neurons in the l_{th} hidden layer. The output of the j^{th} neuron from the l^{th} hidden layer is determined as Equation 27.

$$\emptyset_j^l(x) = f_j^l(.)\left(\sum_{i=l}^{h_l} w_{ji}^l \emptyset_i^{l-1}(x)\right) \quad (27)$$

Where $f_j^l(.)$ is the j^{th} node activation function in the hidden layer l^{th}, the connecting weight w_{ji}^l to the hidden layer j^{th} with i^{th} neuron in the $(l-1)^{th}$ layer. Lian and Douglas (2015) defined the first hidden layer output of j^{th}neuron as given in Equation 28, while the output layer of the k^{th} is given in Equation 29:

$$\emptyset_j^l(x) = f_j^l(.)\left(\sum_{i=l}^{n} w_{ji}^l x_i\right) \quad (28)$$

$$y_k = f_k\left(\sum_{i=l}^{h_l} w_{kj} \emptyset_i^l(x)\right) \quad (29)$$

Where f_k which is the softmax activation function is connected to the output layer and w_{kj} is weight connecting k^{th}output node with the j^{th} node at the last hidden layer.

Training the neural network involved an optimization process using back-propagation algorithm (Hassan et al., 2023; Brochu et al., Defazio and Jelassi, 2022; Assran et al., 2019) as one of the most popular training algorithms for neural network in literature. However, Kayri (2016) revealed that it is slow to convergence and has potential for model over-fitting. Regularization techniques have been proposed to address overfitting (Demuth and Beale, 2000), alongside innovative training algorithms such as Bayesian regularization (Garnett, 2023), advanced RMSprop (Garnett, 2023), Levenberg–Marquardt (Demuth and Beale, 2000), and Adam (Li et al., 2023). However, these methods face limitations. Hassan et al. (2023) reported challenges such as poor hyperparameter space exploration, difficulty due to numerous hyperparameters, unreliability with small datasets, and limited adaptability, which affect their performance in certain machine learning and deep learning applications. To address these challenges and ensure an adaptive fault detection model after training, we propose a combination of Adam and Bayesian algorithm to formulate a hybrid optimization algorithm as shown in Table 1. The Bayesian gradient is a probabilistic model which uses Gaussian process to estimate the model objective function performance and then search for optimal hyper-parameter to minimize loss function. However, Li et al. (2023) and Coleman and Zalewski (2011) revealed that it can get stuck with local optimal and not adaptive. Nevertheless, its fine-tuning capabilities make it suitable in this work to improve Adam optimization algorithm and generate the proposed hybrid optimization algorithm. The new hybrid model for the training process is aimed at improving generalization of model, balance exploration and exploitation of search space, avoid over fitting and improve model scalability. The parameters are learning rate δ, objective function f, $\gamma 1$ is first momentun, $\gamma 2$ second decay momentum, and bias represented as βM_k for the first momentum and βv_k for the second momentum, θ_{k+1}istheweightupdate, $\theta_k$is the state of the model, M_k is the accelerated momentum with exponential weighted gradient, R is the smoothening term, acquisition function αi, while v_k is the is the cumulative square gradient.

Table 1 below presents the proposed hybrid optimization algorithm for the training of the MLP. The proposed hybrid model was developed with the Bayes and training of the neural network was done experimentally using the three different training algorithms. The tool used for the training process is the MATLAB 2022, using neural network toolbox. The process began by coding the proposed algorithms in MATLAB environment, then integrating it with neural network toolbox to train the MLP using the extracted PV fault data in Equation 28. The neural network tool automatically splits the data into training, test and validation set respectively in a ratio of 80:10:10. The training algorithms were carried out experimentally as shown in Figure 5, while the neural network structure and training parameters were reported in Table 2. During the training process, parameters in Equations 30-32 were applied to evaluate the model performance, and when the stopping criteria (tolerable error) were achieved, the model for PV fault prediction was generated. Real data of earth fault, open circuit fault, panel fault, short circuit fault, shadow fault were uploaded individually to evaluate the model performance.

Table 1 Training Algorithms

Adam	Bayesian	Proposed Hybrid optimizer
1) Input: $\theta, m_o = 0, V_o = 0,$ $\gamma 1, \gamma 2 \in [0.1], K = 0, \delta, f,$	1) Input: f, hyper parameter space H, initial data D_o	1) Initialization of parameters
2) While θ not converged,	2) For i=1…….max_ iteration do	2) Set $D_i = D_o$ for the initial data
3) Do;	3) Fit Gaussian process to the data D_i	3) For i=1,……max iteration do
4) $K = k + 1$	4) Maximize$\alpha_i$over H to determine new $x_{i+1} = \text{argmin}_{x \in H} \alpha_i(x)$	4) Fit Gaussian process to the data D_i
5) $M_k = \gamma 1\, M_k - 1 + (1 - \gamma 1)\frac{dl}{d\theta_k}$	5) Evaluate $\text{score}_{i+1} = f(x_{i+1})$	5) Maximize the acquisition function $\alpha_i$over H to determine new $x_{i+1} = \text{argmin}_{x \in H} \alpha_i(x)$
6) $v_k = \gamma 2 v_{k-1} + (1 - \gamma 2)\left[\frac{dl}{d\theta_k}\right]^2$	6) Update$D_{i+1} = D_i(x_i, \text{score}_i)$	6) While θ not converged, D_o;
	7) End for Output (x_{max} score_i)	7) $K = k + 1$ % increase iteration steps

7) $\beta M_k = \frac{M_k}{1-\gamma^k}$
8) $\beta v_k = \frac{v_k}{1-\gamma 1^k}$
9) $\theta_{k+1} = \theta_k - \delta\, \beta M_k / \sqrt{\beta v_k + R}$
10) End while
Output: θ_k

8) $\left.\begin{aligned} M_k &= \gamma 1\, M_k - 1 + (1-\gamma 1)\frac{dl}{d\theta_k} \\ v_k &= \gamma 2 v_{k-1} + (1-\gamma 2)\left[\frac{dl}{d\theta_k}\right]^2 \end{aligned}\right\}$
9) %Update of moment estimates with step 8
10) End

The training algorithms were carried out experimentally as shown in Figure 5, while the neural network structure and training parameters were recorded in Table 2. During the training process, parameters in Equations 30-32 were applied to evaluate the model performance, and when the stopping criteria (tolerable error) were achieved, the model for PV fault prediction was generated. Real data of earth fault, open circuit fault, panel fault, short circuit fault, shadow fault were uploaded individually to evaluate the model performance.
Figure 5 shows the flow chart applied to generate the model for fault detection in the PV system. This model was used to develop the flow chart of the fault management system of Figure 6.

Table 2: Structure of the Neural Network and the Training Parameters

Parameters	Values	Training Parameter	Values
Number of neurons in the input layer	4	Batch size	100
Number of neurons in the first hidden layer	12	Learning rate	0.001
Number of neurons in the second hidden layer	18	Regularization strength	0.0001
Number of neurons in the third hidden layer	32	Epoch size	1000
Activation function in the hidden layers	Tangent hyperbolic	Dropout rate	10%
Number of neurons in the output layer	2	Initial learning rate	0.5
Activation function in the output layer	SoftMax		
Training algorithm	Bayes, Adam, Hybrid		
Loss function	MSE, RMSE		
Regularization	Dropout		

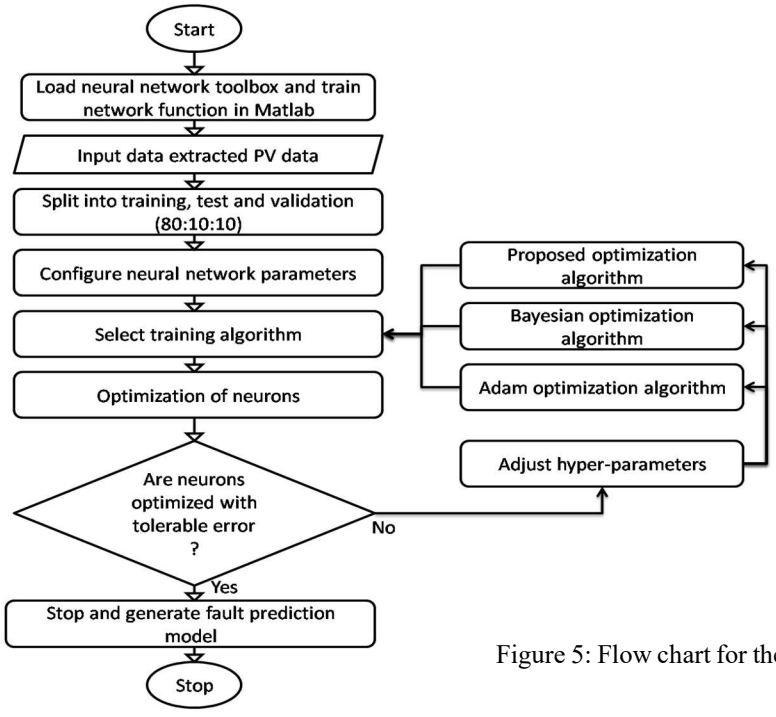


Figure 5: Flow chart for the MLP training

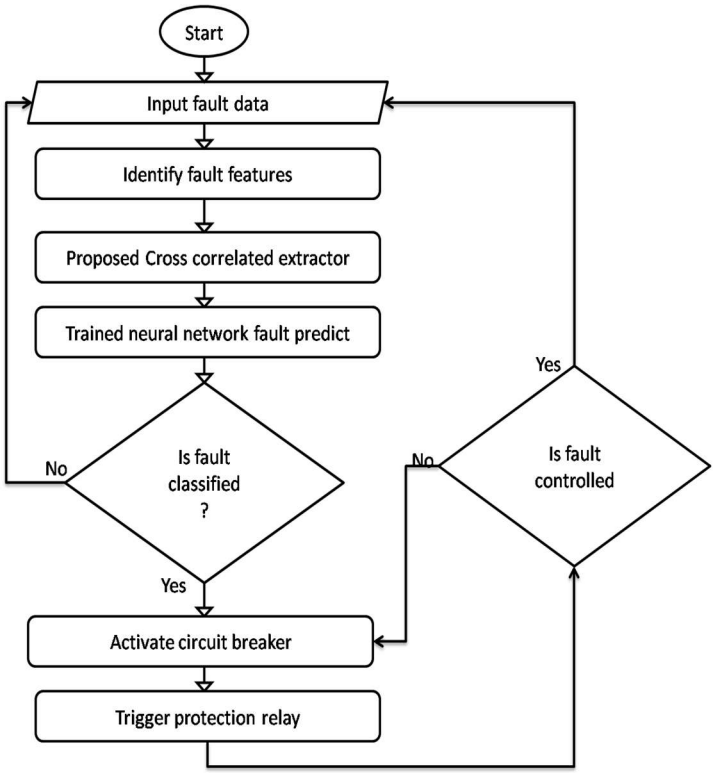


Figure 6 Flow chart fault Management

Figure 6 presents the flow chart for fault management. First the input data from the PV were identified and the features extracted through the proposed cross correlation extractor model and then fed to the trained neural network for prediction of fault. Upon fault detection, circuit breaker is activated which triggers the protection relay to isolate the PV components from fault to ensure real-time protection.

E. Performance Evaluation

Fault detection accuracy can be evaluated using real PV fault data collected from the testbed and then monitor the regression (Habtemariam et al., 2023). The Equation for Mean Square Error (MSE), coefficient of determination R^2, Mean Absolute Error (MAE),

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y - \breve{y}| \quad (30)$$

$$MSE = \sum_{i=1}^{n}\frac{(y-\breve{y})^2}{N} \quad (31)$$

$$R^2 = 1 - \frac{\frac{n}{i=1}(y_i - \widehat{y_i})^2}{\frac{n}{i=1}(y_i - \breve{y})^2} \quad (32)$$

Where y is the fault dataset extracted with cross correlations, $\breve{y}$ is the predicted fault data, N is the overall observation applied to train the model, and $\widehat{y_I}$ is the mean actual values.

IV. RESULTS AND DISCUSSIONS

The results of the neural network training were discussed in this section considering the RMSE, MSE, and R^2respectively as reported. Figure 7-9 presents the MSE results for the three respective training algorithms during validation with real data of PV fault. Figure 7 shows the MSE with Adam optimization, reporting MSE of 9.6795e^{-11}, while the MSE with Bayes reported 2.2285e^{-11} in Figure 8 and MSE with hybrid optimization algorithm reported 9.3103e^{-14} in Figure 9. Table 3 summarizes the performance of the trained neural network. Figure 6: Flow chat for fault management system

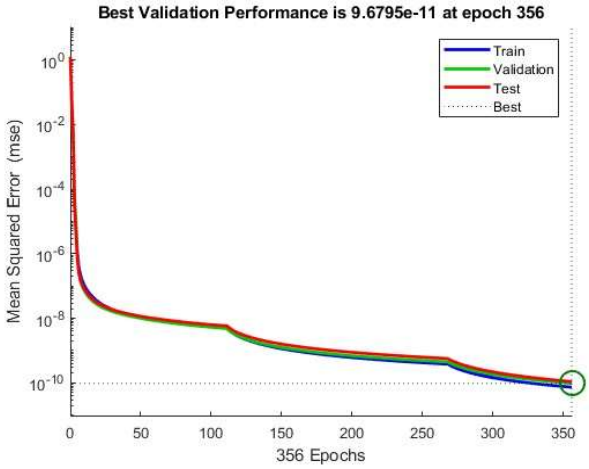


Figure 7 MSE with Adam

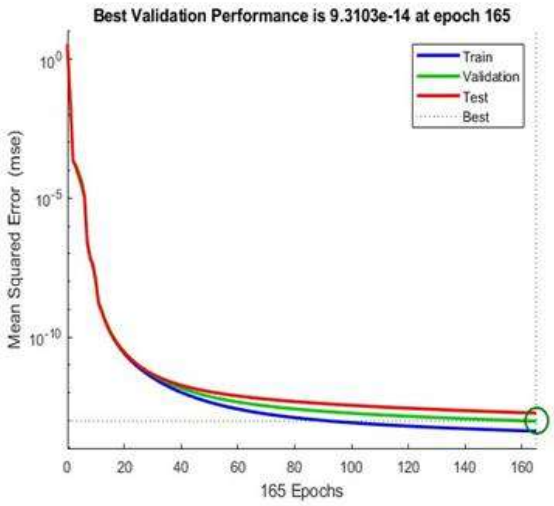


Figure 8: MSE with Hybrid algorithm

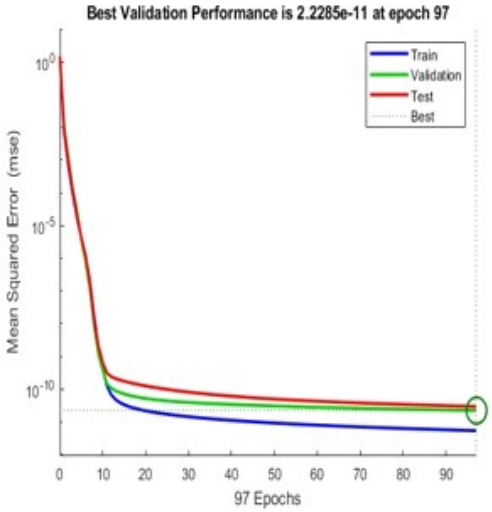


Figure 9: MSE with Bayesian

Table 3 reports the performance of the PV fault detection model generated for early fault management. The results show the three different training algorithms employed to train the MLP and their respective performances. The result showed that overall, the three-training algorithm recorded MSE values which are acceptable as they all indicated minimal loss. The RMSE also recorded similar results for the three different algorithms, implying minimal error when the actual value was matched with the predicted values of PV fault data. The R^2 also indicated the success of fitness between the true PV fault data and the predicted PV fault data. Overall, the results for the Bayesian optimizer are (2.2285e-11 for MSE, 4.72069e^{-6} for RMSE and 0.9704 for R^2). Adam reported (9.6795e^{-11} for MSE, 9.838445e^{-6} for RMSE and 0.9789 for R), while the proposed hybrid optimization algorithm reported 9.3103e^{-14} for MSE, 3.051278e^{-7} for RMSE and 0.9894 for R^2. After training the model, several test data which was obtained during fault condition of the testbed were loaded into the neural network tool to test the model again. The results were recorded in Tables 4-7.

Table 3: Final Training Result of the PV Fault Detection Model

Parameters	Bayesian	Adam	Hybrid optimization
MSE	2.2285e^{-11}	9.6795e^{-11}	9.3103e^{-14}
RMSE	4.72069e^{-6}	9.838445e^{-6}	3.051278e^{-7}
R^2	0.9704	0.9789	0.9894

Table 4 Shadow Fault

Parameters	**Bayesian**	**Adam**	**Hybrid Optimization**
MSE	3.6458e^{-9}	9.6435e^{-9}	11.3103e^{-13}
RMSE	6.0380e^{-5}	9.8201e^{-5}	1.06343e^{-6}
R^2	0.9573	0.9664	0.9858

Table 5 Short Circuit Fault

Parameters	**Bayesian**	**Adam**	**Hybrid optimization**
MSE	2.2285e^{-7}	9.6795e^{-7}	9.3103e^{-11}
RMSE	4.7206e^{-6}	9.8384e^{-6}	3.05129e^{-7}
R^2	0.95333	0.9537	0.9743

Table 6 Earth fault

Parameters	**Bayesian**	**Adam**	**Hybrid optimization**
MSE	2.7653e^{-5}	9.8832e^{-5}	7.6563e^{-8}
RMSE	5.2586e^{-6}	9.3978e^{-6}	2.76703e^{-7}
R^2	0.9623	0.9657	0.9854

Table 7 Open circuit fault

Parameters	**Bayesian**	**Adam**	**Hybrid optimization**
MSE	3.654e^{-5}	9.7545e^{-7}	9.5673e^{-9}
RMSE	0.00614	9.8766e^{-4}	9.7818e^{-5}
R^2	0.9756	0.9635	0.9824

Table 4-7 records the test results of the trained model with fault data collected from the testbed. The results in Table 4 reported the MSE, R^2 and RMSE when tested with shadow fault. In these results, the data obtained when part of the panel was covered was applied to test the model. The results show that for Bayes (MSE is 3.6458e^{-9}, RMSE is 6.0380e^{-5} and R^2 is 0.9573), for Adam the MSE is 9.6435e^{-9}, RMSE is 9.8201e^{-5} and R^2 is 0.9664, while hybrid optimization algorithm reports MSE of 11.3103e^{-13}, RMSE of 1.06343e^{-6} and R^2 of 0.9858 respectively. The model prediction performance for short circuit fault in Table 5 considering MSE with Bayes reported 2.2285e^{-7}, RMSE of 4.7206e^{-6} and R^2 of 0.95333. When tested with Adam generated model the RMSE reported 9.8384e^{-6}, MSE is 9.6795e^{-7} and R^2 is 0.9537, while with the hybrid algorithm reported MSE of 9.3103e^{-11}, RMSE of 3.05129e^{-7} and R^2 of 0.9743. Table 6 reported results of earth fault with Bayes for MSE as 2.7653e^{-5}, RMSE as 5.2586e^{-6} and R^2 as 0.9623. Adam optimization algorithm reported MSE of 9.8832e^{-5}, RMSE of 9.3978e^{-6} and R^2 of 0.9657.

Hybrid optimization algorithm when tested with earth fault reported MSE of 7.6563e^{-8}, RMSE of 2.76703e^{-7} and R^2 of 0.9854. Table 7 reported the testing result for Bayesian MSE at 3.654e^{-5}, RMSE at 0.00614 and R^2 of 0.9756. Adam reported MSE of 9.7545e^{-7}, RMSE of 9.8766e^{-4} and R^2 of 0.9635, while Hybrid optimization reported MSE of 9.5673e^{-9}, RMSE of 9.7818e^{-5} and R^2 of 0.9824. The percentage reduction in MSE with hybrid optimization is 99.96%. The absolute difference in the R^2 considering open circuit fault with Adam against hybrid is 0.0189 while percentage improvement is 1.96%. For RMSE considering earth fault, the hybrid optimization recorded 19.01 times less error compared with Bayesian, while the percentage reduction in error for open circuit fault with hybrid optimization against Adam reported 90.94%. These results has consistently shown that the hybrid optimization training algorithm recorded less error and better improvement in correctly predicting different classes of faults in PV system. The average overall R^2 for Bayesian is 0.9621; Adam is 0.9625 and hybrid 0.9820, the average percentage improvement with hybrid recorded 2.07% against Bayesian and 2.03% against Adam.

V. CONCLUSION

Sustainable energy takes more than just the availability of power but also considers reliability and efficiency. Over time, fault has remained a major challenge to achieve this goal of sustainable power supply with PV, and the easy approach to manage the problem is through early detection and control. This paper has trained MLP neural network to generate a fault classification model which was applied to develop a model for the detection and management of fault in 4.2kW PV power system. This was achieved through data collection of different PV fault scenarios. The data were extracted using cross correlation approach and then applied to train a MLP neural network using Bayesian, Adam and proposed hybrid optimization algorithms. Three fault prediction models were generated and compared across different fault sample data. The results obtained showed hybrid optimization algorithm produced the most suitable model for PV fault prediction.

AUTHORS CONTRIBUTIONS

C. U. Nwamuo: Conceptualization, Methodology, Software, Validation, Writing-original draft, Writing-review and editing. **O. Ojike**: Supervision, Editing. **S. O. Enibe**: Supervision, Editing.

ACKNOWLEDGEMENT

The researchers wish to extend their sincere gratitude to the African Centre of Excellence for Sustainable Power and Energy Development and University of Nigeria Nsukka for the opportunity provided to study this program.

REFERENCES

Adrian-Martinez, S., & Ardid, M., (2015). Acoustic Signal detection Through the Cross-correlation Methods in Experiments with Different Signal to Noise Ratio and Reverberation Conditions. *Journal of Instrumentation and Detectors*.

Al-Katheri, A.A., Al-Ammar, E.A., Alotaibi, M.A., Ko, W., Park, S., & Choi, H.-J. (2022). Application of Artificial Intelligence in PV Fault Detection. *Sustainability* **2022**, *14*, 13815. https://doi.org/10.3390/su142113815

Appiah, A.Y., Zhang, X., Ayawli, B.B.K., & Kyeremeh, F. (2019). Review and Performance Evaluation of Photovoltaic Array Fault Detection and Diagnosis Techniques. *Int. J. Photoenergy* **2019**, *2019*, 6953530.

Assran, M., Loizou, N., Ballas, N., & Rabbat, M. (2019). Stochastic gradient push for distributed deep learning. 36th Int Conf Mach learn ICML 2019 2019-June:514–523. https://proceedings.mlr.press/v97/assran19a.html

Begum, M., Shorif, S.B., Uddin, M.S., Ferdush, J., Jan, T., Barros, A., & Whaiduzzaman, M. (2024). Image Watermarking Using Discrete Wavelet Transform and Singular Value Decomposition for Enhanced Imperceptibility and Robustness. *Algorithms 17*, 32. https://doi.org/10.3390/a17010032

Bishop, C.M. (1995). Neural Networks for Pattern Recognition: *Oxford University Press*.

Bose, R., Samanta, K., & Chatterjee, S. (2016). Cross-correlation based feature extraction from EMG signals for classification of neuro-muscular diseases. *International Conference on Intelligent Control Power and Instrumentation (ICICPI)*, Kolkata, India, 2016, pp. 241-245, doi: 10.1109/ICICPI.2016.7859710.

Brochu, E., Cora, V.M., & De Freitas N (2010). A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcementlearning.arXivPreprarXiv10122599. https://doi.org/10.48550/arXiv.1012.2599

Camacho, N. J., Ruiz, M., Villamizar, R., Mujica, L., & Quiroga, J. (2018). Features of Cross-Correlation Analysis in a Data-Driven Approach for Structural Damage Assessment. *Sensors* **2018**, *18*, 1571. https://doi.org/10.3390/s18051571

Chen, D., Wan, S., & Bao, F.S. (2017). Epileptic Focus Localization Using Discrete Wavelet Transform Based on Interictal Intracranial EEG. *IEEE Trans Neural Syst Rehabil Eng*. 25(5):413-425. doi:

10.1109/TNSRE.2016.2604393. Epub 2016 Aug 30. PMID: 28113594.

Chen, L., Han, W., & Cao, X. (2020). Long distance wireless fault diagnosis for photovoltaic modules based on back propagation neural network. *Int. Journal of Electrical EngEducat*. https://doi.org/10.1177/0020720920940601

Chen, Z., Chen, Y., Wu, L., Cheng, S., & Lin, P. (2019). Deep Residual Network Based Fault Detection and Diagnosis of Photovoltaic Arrays Using Current-Voltage Curves and Ambient Conditions. *Energy Convers. Manag*. **2019**, *198*, 111793.

Chidi, E. U., Udanor, C. N., & Anoliefo, E. (2024). Exploring the Depths of Visual Understanding: A Comprehensive Review on Real-Time Object of Interest Detection Techniques. *Preprints* **2024**, 2024020583. https://doi.org/10.20944/preprints202402.0583.v1

Chine,W., Mellit, A., Lughi, V., Malek, A., Sulligoi, G., & Pavan, A.M. (2016). A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. *Renew. Energy* 90, 501–512.

Coleman, A., & Zalewski, J. (2011). Intelligent Fault Detection and Diagnostics. *In Proceedings of the The 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems:* Technology and Applications, Prague, Czech Republic, 15–17 September 2011

Dau, A., & Salim, N. (2020). Recommendation system based on deep learning methods: a systematic review and new directions. *Artif Intel Rev*. 2020;53(4):2709–48.

Defazio, A., & Jelassi, S. (2022). Adaptivity without compromise: a momentumized, adaptive, dual averaged gradient method for stochastic optimization. J Mach Learn Res 23:1–34

Demuth, H., & Beale, M. (2000). *Neural Network Toolbox User's Guide Version 4*; The Math Works Inc.: Natick, MA, USA, pp. 5–22.

Deng, L. (2014) A tutorial survey of architectures, algorithms, and applications for deep learning. APSIPA Trans Signal Inf Process. 2014; p. 3.

Ebere, U., Anoliefo, E., Udanor, C., Chijindu, T., & Nwobodo, L. (2025). A Blind navigation guide model for obstacle avoidance using distance vision estimation based YOLO-V8n; Journal of the Nigerian Society of Physical Sciences, 2292-229; https://doi.org/10.46481/jnsps.2025.2292

El-Banby, M., Moawad, M., & Abouzalm, A. (2023). Photovoltaic system fault detection techniques: a review. *Neural Computing&Application* **35**, 24829–24842. https://doi.org/10.1007/s00521-023-09041-7

El-Rashidy, M.A. (2022). An efficient and portable solar cell defect detection system. Neural Comput Appl 34:18497–18509. https:// doi.org/10.1007/s00521-022-07464-2

Et-Taleby, A., Chaibi, Y., Benslimane, M., & Boussetta M., (2023) Applications of Machine Learning Algorithms for Photovoltaic Fault Detection: a Review. Statistics, Optimization and Information Computing Stat., Optim. Inf. Comput., Vol. 11, January 2023, pp 168–177. DOI: 10.19139/soic-2310-5070-1537

Firth, S., Lomas, K., & Rees, S. A. (2010). Simple model of PV system performance and its use in fault detection. *Sol. Energy***2010**, 84, 624–635.

Gao, W., & Wai, R.J. (2020). A Novel Fault Identification Method for Photovoltaic Array via Convolutional Neural Network and Residual Gated Recurrent Unit. *IEEE Access* **2020**, *8*, 159493–159510

Garnett, R. (2023). Bayesian Optimization; Cambridge University Press: Cambridge, UK.

Garoudja, E., Harrou, F., Sun, Y., Kara, K., Chouder, A., & Silvestre, S. (2017). Statistical fault detection in photovoltaic systems. *Solar Energy* **2017**, *150*, 485–499.

Global Cumulative Installed Solar PV Capacity 2019|Statista. Available online: https://www.statista.com/statistics/280220 /global-cumulative-installed-solar-pv-capacity (accessed on 10 September 2024).

Habtemariam, E.T., Kekeba, K., Martínez-Ballesteros, M., & Martínez-Álvarez, F. (2023). A Bayesian Optimization-Based LSTM Model for Wind Power Forecasting in the Adama District, Ethiopia. *Energies 16*, 2317. https://doi.org/10.3390/en16052317

Harbor, M.C., Eneh, I.I., & Ebere U.C. (2021). Nonlinear dynamic control of autonomous vehicle under slip using improved back-propagation algorithm. International Journal of Research and Innovation in Applied Science (IJRIAS); Vol. 6; Issue 9; https://rsisinternational.org/journals/ijrias/DigitalLibrary/volume-6-issue-9/62-68.pdf

Hassan, E., Shams, M.Y., & Hikal, N.A. (2023). The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study. *Multimed Tools Appl* **82**, 16591–16633. https://doi.org/10.1007/s11042-022-13820-0

Heinrich Haberlin, (2016). ''Photovoltaic system design and practice'' second Edition , John Wiley & Sons, Ltd.

Hong, Y.Y., & Pula, R.A. (2022). Methods of Photovoltaic Fault Detection and Classification: A Review. *Energy Rep.* **2022**, *8*, 5898–5929.

Ibrahim, B. M., Taha Nagy, I., Alsharef, M., Sabiha, N. A., Metwaly, M. K., & Elattar, E. E. (2020). Investigation of diode dynamic effect on fault detection of photovoltaic systems. *Sustainable Energy, Grids and Networks, 23*, 100389. https://doi.org/10.1016/j.segan.2020.100389

Kachhwaha, A., Rashed, G. I., Garg, A. R., Mahela, O. P., Khan, B., Shafik, M. B., & Hussien, M. G. (2022). Design and Performance Analysis of Hybrid Battery and Ultracapacitor Energy Storage System for Electrical Vehicle Active Power Management. *Sustainability*, *14*(2), 776.

Kayri, M. (2016). Predictive Abilities of Bayesian Regularization and Levenberg–Marquardt Algorithms in Artificial Neural Networks: A Comparative Empirical Study on Social Data. *Math. Comput. Appl. 21*, 20. https://doi.org/10.3390/mca21020020

Kekong, P.E, Ajah, I.A., & Ebere, U.C. (2019). Real-time drowsy driver monitoring and detection system using deep learning based behavioural approach. International Journal of Computer Sciences and

Engineering 9 (1), 11-21; http://www.ijcseonline.isroset.org/pub_paper/2-IJCSE-08441-18.pdf

Khelil, C.K.M., Amrouche, B., Kara, K., & Chouder, A. (2021). The impact of the ANN's choice on PV systems diagnosis quality. *Energy Convers. Manag*. 240, 114278.

Kontges, M., Kurtz, S., Packard, C., Jahn, U., Berger, K., Kato, K., Friesen, T., Liu, H., Iseghem, M.V., Wohlgemuth, J., Miller, D., Kempe, M., Hacke, P., Reil, F., Bogdansk, N., Herrmann, W., Buerhop-Lutz, C., Razongles, G., & Friesen G (2014). Review of failures of photovoltaic modules—*iea-pvps* [www document]. https://iea-pvps.org

Lakshmi, P., Sivagamasundari, S., & Rayudu, M. (2023) IoT based solar panel fault and maintenance detection using decision tree with light gradient boosting. Measurement: *Sensors* 27 (2023) 100726 https://doi.org/10.1016/j.measen.2023.100726

Lee, H., & Yun, S. (2024). Strategies for Imputing Missing Values and Removing Outliers in the Dataset for Machine Learning-Based Construction Cost Prediction. *Buildings 14*, 933. https://doi.org/10.3390/buildings14040933

Lee, S. (2020). Performance Comparison of Imputation Algorithms for Missing Data. Master's Thesis, Graduate School of Korea University, Seoul, Republic of Korea.

Lee, S., & Jeon, S. (2009). An Empirical Comparison of Imputation Methods for Missing Data. *KIIS Spring Conf. 19*, 250–251.

Li, B., Delpha, C., Diallo, D., & Migan-Dubois A. (2019). Application of Artificial Neural Networks to photovoltaic fault detection and diagnosis: A review. *Renewable and Sustainable Energy Reviews*https://doi.org/10.1016/j.rser.2020.110512

Li, Z., Pan, Z., Li, Y., Yang, X., Geng, C., & Li, X. (2023). Advanced root mean square propagation with the warm-up algorithm for fiber coupling. *Opt. Express*, 31, 23974–23989.

Lian, J., & Douglas, L. (2015). Automatic Fault Detection and Diagnosis for Photovoltaic Systems using Combined Artificial Neural Network and Analytical Based Methods"; 978-1-4799-1959-8/15/ *IEEE*

Liu, H., Perera, A., Al-Naji, A., Boubaker, S., Kamel, S., Ghazouani, N., & Mellit, A. (2023). Assessment of Machine and Deep Learning Approaches for Fault Diagnosis in Photovoltaic Systems Using Infrared Thermography. *Remote Sens.* **2023**, *15*, 1686

Liu, Y., Ding, K., Zhang, J., Li, Y., Yang, Z., Zheng, W., & Chen, X. (2021). Fault Diagnosis Approach for Photovoltaic Array Based on the Stacked Auto-Encoder and Clustering with I-V Curves. *Energy Convers. Manag.* **2021**, *245*, 114603.

Livera, A., Theristis, M., Makrides, G., & Georghiou, G.E. (2019). Recent advances in failure diagnosis techniques based on performance data analysis for grid-connected photovoltaic systems.*Renewable Energy*, vol.133, no. 11, pp. 126–143.

Mekki, H., Mellit, A., & Salhi, H. (2016). Artificial neural network-based modelling and fault detection of partial shaded photo-voltaic modules. Simul. Model. Pract. Theory 67, 1–13.

Mellit, A., Tina, G.M., & Kalogirou, S.A. (2018). Fault detection and diagnosis methods for photovoltaic systems: a review. *Renew Sustain Energy Rev* 91:1–17

Mellit, A., Benghanem, M., Kalogirou, S., & Pavan A. (2023). An embedded system for remote monitoring and fault diagnosis of photovoltaic arrays using machine learning and the internet of things. Renewable Energy 208 (2023) 399–408 https://doi.org/10.1016/j.renene.2023.03.096

Mouaad, B., Samir, B., & Mohand, A. (2024). Dynamic Modeling for Fault Diagnosis in PV Systems Utilizing AI Techniques Based on Multilayer Perceptron. Conference: 2024 *international conference in emerging technologies for sustainability and intelligent systems (icetsis 2024); doi*: 10.1109/ICETSIS61505.2024.10459490

Rafique, H., & Rafique, S. (2016). Review of cross-correlated based algorithm or image and signal processing or pattern identification. *International journal of Geomate*; Vol 11; Issue 27; pp. 2695-2703.

Rao, S., Spanias, A., & Tepedelenlioglu, C. (2019). Solar array fault detection using neural networks. *In Proceedings of the 2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS),* Taipei, Taiwan, 6–9 May 2019; pp. 196–200.

Rouani, L., Harkat, M., Kouadri, A., & Mekhilef S. (2021). Shading fault detection in a grid-connected PV system using vertices principal component analysis. *Renewable Energy* 164 (2021) 1527e1539 https://doi.org/10.1016/j.renene.2020.10.059

Sairam, S., Seshadhri, S., Marafioti, G., Srinivasan, S., Mathisen, G., & Bekiroglu K., (2022). Edge-based Explainable Fault Detection Systems for photovoltaic panels on edge nodes. Renewable Energy 185 (2022) 1425e1440 https://doi.org/10.1016/j.renene.2021.10.063

Sarker, I.H. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN COMPUT. SCI.* **2**, 420 (2021). https://doi.org/10.1007/s42979-021-00815-1

Stähler, S.C., Sens-Schönfelder, C., & Niederleithinger, E. (2011). Monitoring stress changes in a concrete bridge with coda wave interferometry.*J. Acoust. Soc. Am.*, vol. 129, no. 4, pp. 1945–1952, 2011., http://doi.org/10.1121/1.3553226.

Tina, G.M., Cosentino, F., &Ventura, C. (2016). Monitoring and diagnostics of photovoltaic power plants. In *Renewable Energy in the Service of Mankind Volume II*; Springer: Berlin, Germany, 2016; pp. 505–516.

Ul-Haq, A., Sindi, H.F., Gul, S., & Jalal, M. (2020). Modeling and fault categorization in thin-film and crystalline PV arrays through multilayer neural network algorithm. *IEEE Access* 8, 102235–102255.

Utama, C., Meske, C., Schneider, J., Schlatmann, R., & Ulbrich, C. (2023). Explainable artificial intelligence for photovoltaic fault detection: A comparison of instruments. Solar Energy 249 (2023) 139–151https://doi.org/10.1016/j.solener.2022.11.018

Wang, P., Li, C., Liang, R., Yoon, S., Mu, S., & Liu, Y. (2023). Fault detection and calibration for building energy system using Bayesian inference and sparse autoencoder: A case study in photovoltaic thermal heat

pump system. Energy & Buildings 290 (2023) 113051 https://doi.org/10.1016/j.enbuild.2023.113051

Wang, W., Liu, A. C. -F., Chung, H. S. -H., R. Lau, W. -H., Zhang, J., & Lo, A. W. -L. (2016). Fault Diagnosis of Photovoltaic Panels Using Dynamic Current–Voltage Characteristics. *IEEE Transactions on Power Electronics, vol. 31, no. 2, pp. 1588-1599, Feb. 2016,* doi: 10.1109/TPEL.2015.2424079.

Weidong X. (2017). ‘‘Photovoltaic Power System Modelling, Design, and Control’’, 1st, Edition, John Wiley& Sons Ltd.

Yuan, Z., Xiong, G., & Fu, X. (2022). Artificial Neural Network for Fault Diagnosis of Solar Photovoltaic Systems: A Survey. Energies **2022**, 15, 8693. https://doi.org/10.3390/ en15228693

Zhang, Y., Abraham, O., Grondin, F., Loukili, A., Tournat, V., Le Duff, A., Lascoup, B., & Durand, O. (2012). Study of stress-induced velocity variation in concrete under direct tensile force and monitoring of the damage level by using thermally-compensated coda wave interferometry. *Ultrasonics*, vol. 52, no. 8, pp. 1038–1045, 2012., http://dx.doi.org/10.1016/j.ultras.2012.08.011.

Zhang, Z., Ma, M., Ma, W., Zhang, R., & Wang, J. (2022). A data-driven photovoltaic string current mismatch fault diagnosis method based on I-V curve. Microelectronics Reliability 138 (2022) 114705 https://doi.org/10.1016/j.microrel.2022.114705