



الجمهورية الجزائرية الديمقراطية الشعبية
People's Democratic Republic of Algeria

وزارة التعليم العالي والبحث العلمي
Ministry of Higher Education and Scientific Research

المدرسة الوطنية العليا للتكنولوجيات المتقدمة
National Higher School of Advanced Technologies



Department of Electrical Engineering and Industrial Computer Science

Final Year Project to Obtain the Diploma

Field : **Electronics**

Specialty : **Embedded Systems**

Fault Detection and Diagnosis of PV Systems Based on Corrected Performance Ratio Analysis and Adaptive Thresholds Using CNN-LSTM Hybrid Neural Network Prediction Model

Realized by :

Abderrahim ADDALA

Chahine KHALED

Full Name	University	Grade	Quality
Ms. Karima BOULAAM	ENSTA	MCB	President
Ms. Lamia HABANI	ENSTA	MAA	Examinator
Mr. Mohamed Lassaad BENZAOUI	ENSTA	MAA	Examinator
Mr. Reda DERMOUCHE	ENSTA	MCA	Supervisor
Mr. Aissa CHOUDER	U.Msila	Professor	Supervisor
Ms. Imane DROUCHE	ENSTA	MCB	Supervisor

Academic year : 2024-2025

Fault Detection and Diagnosis of PV Systems Based on Corrected Performance Ratio analysis and Adaptive Thresholds Using CNN-LSTM Hybrid Neural Network Prediction Model

Abderrahim ADDALA
National School of Advanced
Technologies
Alger, Algeria
a_addala@ensta.edu.dz

Chahine KHALED
National School of Advanced
Technologies
Alger, Algeria
c_khaled@ensta.edu.dz

Aissa CHOUDER
University of M'Sila
M'Sila, Algeria
aissachouder@gmail.com

Reda DERMOUCH
National School of Advanced
Technologies
Alger, Algeria
reda.dermouche@ensta.edu.dz

Oudira HOUCINE
University of M'Sila
M'Sila, Algeria
houcine.oudira@univ-msila.dz

Imane DROUCHE
National School of Advanced
Technologies
Alger, Algeria
imane.drouiche@enst.dz

Abstract—In the model-based approach, a trusted and reliable photovoltaic (PV) array model provides an essential baseline for monitoring the system's health and diagnosing any underperformance. To this end, this article proposes a three-phase methodology for developing a robust PV plant model and implementing a fault detection and diagnosis mechanism, suitable for large-scale grid-connected PV systems. The first stage involves constructing a single-layer hybrid CNN-LSTM model aimed at predicting photovoltaic power generation. This model uses real-world measured data obtained from a large-scale, grid-connected photovoltaic system operating in Algeria. In this architecture, the convolution layer acts as a feature extractor, identifying local patterns in the measured data. Meanwhile, the LSTM component captures temporal dependencies. In the second phase, an automated inspection platform is developed, leveraging the evaluation of the Corrected Performance Ratio (CPR) for each inverter within the operational photovoltaic (PV) system... Next, a lower threshold and an upper threshold set by Hampel identifier are defined and processed, allowing the system to detect faults by generating a faulty signal. The diagnostic process is implemented by setting dynamic voltage and current thresholds derived from the prediction model. These thresholds are then adjusted adaptively to improve the sensitivity of fault detection for both voltage and current abnormalities. In the final phase, the proposed supervision system is implemented as a standalone graphical user interface (GUI) application based on the developed method-

ology. This validates its applicability for photovoltaic (PV) supervisory control and data acquisition (SCADA) systems. Evaluated case studies demonstrate the effectiveness of this approach in detecting typical PV systems faults, such as short circuits, string disconnections, and partial shading effects.

Keywords—PV prediction; computational modeling; regression techniques, Correction performance ratio, GUI application

1. INTRODUCTION

With the global push towards renewable energy, photovoltaic systems have emerged as essential components of sustainable power generation [1][2][3][4]. However, their performance and efficiency are regularly hampered by several faults, leading to significant power losses, poor performance, and reduced output. Over time, these issues result in high operating costs due to increased maintenance needs. To enhance the reliability and sustainability of PV systems, effective fault detection systems are necessary. Despite advances in PV technology, fault detection in real-world conditions remains challenging due to environmental variability and the diverse nature of faults, including sensor malfunctions [5]. Fault detection methods often struggle with the complexity and scale of large PV installations, particularly when it comes to handling thousands of modules operating in constantly changing weather conditions.

A key step in developing a robust fault detection strategy is establishing an accurate PV array model, which serves as the foundation for identifying and analyzing potential faults, ensuring optimal system performance [6]. This modeling approach faces significant challenges, particularly in predicting PV system output accurately, as PV performance is highly sensitive to the dynamic meteorological factors such as solar

irradiance, wind speed, temperature, cloud cover, and module temperature [7]. These factors exhibit temporal variability, complicating the prediction of energy yield. Thus, predicting PV output is essential not only for fault detection but also for grid integration and stable power supply. To address these challenges, various prediction methods have been developed, including physically based models, data-driven statistical models, and hybrid systems that combine both approaches. Physically based models [8, 9] rely on deterministic equations and meteorological variables but struggle during periods of rapid environmental change. Data-driven models, on the other hand, analyze historical data to detect patterns and provide flexibility in diverse scenarios [10, 11, 12]. In addition to model-based and data-driven prediction techniques, other research has highlighted the use of the Corrected Performance Ratio (CPR) for fault detection in PV systems, which adjusts the conventional performance ratio by accounting for environmental variations to improve fault isolation accuracy. A fault diagnosis framework was proposed using SCADA data from a 40 MW PV plant, incorporating CPR for fault isolation alongside statistical detection and anomaly-based classification. However, its diagnosis phase considers only two fault categories: recoverable and unrecoverable faults [13]. Hybrid systems aim to enhance prediction accuracy under varying conditions [14]. Machine learning (ML) techniques, such as artificial neural networks (ANN), long short-term memory (LSTM) networks, and support vector machines (SVM), have gained prominence in PV power forecasting due to their ability to model complex, non-linear relationships.

While accurate prediction of PV output is crucial, it must be complemented by effective fault detection mechanisms to ensure reliable system operation. Choosing suitable fault detection methods is also challenging, as the effectiveness of these methods directly impacts the timely identification of faults, which is essential for minimizing losses [14]. Traditional threshold-based methods [15, 16], which analyze electrical parameters such as current and voltage, are limited by the need for parameter calibration, which can be resource-intensive and difficult to scale. Machine learning-based methods have been extensively explored to address these challenges. For instance, neural networks have been employed for fault diagnosis by analyzing attributes such as current and voltage under varying operating conditions [17]. Additionally, Liu et al. proposed a fault diagnosis method for photovoltaic arrays utilizing a stacked auto-encoder (SAE) combined with clustering. This approach automatically extracts features by analyzing inherent I-V characteristics, facilitating effective fault detection and classification diagnosis [18]. Similarly, based on output I-V characteristic curves and input ambient condition data, a novel deep residual network (ResNet) based on an intelligent fault detection and diagnosis approach was proposed by Chen et al. [19] to improve diagnostic accuracy. However, these methods often require large, diverse datasets and are sensitive to environmental variability. Recent studies, such as those using convolution neural network (CNN) [20], have achieved high accuracy in identifying anomalies, but their reliance on image-based data and the challenges of generalizing across different environments remain limitations. Moreover, most of these approaches fail to specifically tackle sensor-related faults, such as misaligned pyranometers, which can result in erroneous irradiance readings and ultimately reduce the reliability of fault detection.

To overcome these challenges, the present work's contributions involve developing a robust CNN-LSTM PV model that is the foundation for monitoring and fault detection. Additionally, a fault detection procedure based on CPR analysis is introduced. This method focuses on identifying a lower threshold, determined using the Hampel identifier, to effectively detect and diagnose faults within the PV system. Finally, integrating the supervision system based on the best-performing model into a standalone GUI application based on Python to facilitate predictions. This methodology provides an efficient, scalable solution for detecting and classifying faults in large PV systems, ultimately improving system reliability and reducing maintenance costs.

The adopted methodology unfolds in three crucial phases:

The first phase involves building a reliable and accurate photovoltaic (PV) power prediction model. This model serves as the cornerstone for monitoring the performance of the PV system and identifying anomalies. Historical operational data from the PV plant, including meteorological and electrical parameters (such as irradiance, temperature, P_{DC} , V_{DC} , and I_{DC}), are collected and preprocessed. The model is trained using advanced deep learning architectures, particularly a one-layer CNN-LSTM hybrid, that captures both spatial patterns and temporal dynamics of the data. This model not only predicts expected PV output under normal conditions but also forms a baseline for detecting deviations that may indicate faults. Moreover, this phase includes a correction mechanism for sensor-related anomalies, such as misaligned pyranometers, that could distort the accuracy of environmental inputs. By improving the reliability of input data, the system enhances prediction precision and reduces the risk of false positives or missed detections.

In the second phase, the methodology introduces a fault detection framework that analyzes the Corrected Performance Ratio (CPR) of each inverter in the PV system. A lower threshold, determined using the Hampel identifier, is established to detect faults by generating a fault signal. Additionally, two indicators based on adaptive boundaries derived from predicted current and voltage values are introduced. By evaluating the fault signal against these adaptive limits, the system can accurately identify the fault type (e.g., shading, soiling, degradation, or string disconnection).

The final phase of this paper involves deploying the trained model within a standalone GUI monitoring platform. Each day, the application processes the PV data to calculate key parameters such as yield, energy losses, and the Corrected Performance Ratio (CPR) and then runs the detection and diagnosis procedure using these dynamic metrics. With a scalable backend infrastructure and interactive user interfaces, the platform generates daily fault alerts based on CPR deviations and dynamic prediction boundaries. It also provides end-of-day performance reports and decision-making tools, ultimately enhancing system reliability and reducing operational costs.

The rest of the paper is organized as follows: Section 2 presents an overview of the studied site and PV power plant. Section 3 presents the work methodology, including the predictive models, the fault detection and diagnosis method based on the CPR metric, and dynamic predictive boundaries. Finally, Sections 4 and 5 discuss the results and present the conclusions of the study.

2. EXPERIMENTAL SETUP

2.1 the Dataset

The dataset utilized in this research comes from a grid-connected, ground-mounted photovoltaic (PV) system situated in Ain El-Melh, a region in the Algerian highlands bordering the vast desert. The plant is located at 34°51 N latitude and 04°11 E longitude, with an altitude of 910 meters above sea level. This PV system is connected to Ain El-Melh’s medium-voltage network and forms part of a significant 400 MW renewable energy project led by SKTM, a Sonelgaz subsidiary. Under the mandate of the Algerian government to promote renewable energy, Sonelgaz has established 23 PV power plants across the highlands and central regions. Covering an area of 40 hectares, the Ain El-Melh facility has an installed capacity of 20 MWp and incorporates design features to maximize energy output. The system uses poly-crystalline silicon modules with an efficiency of 15%. These modules are installed on fixed structures at a tilt angle of 33° facing south to maximize solar exposure. The rows of PV modules are spaced 5 meters apart to minimize shading and enhance energy capture. The power plant employs inverters with a nominal power of 500 kW and transformers rated at 1250 kVA operating at 47–52 Hz and 315 V/31.5 kV, ensuring efficient energy conversion and integration into the grid (**Figure 1**).

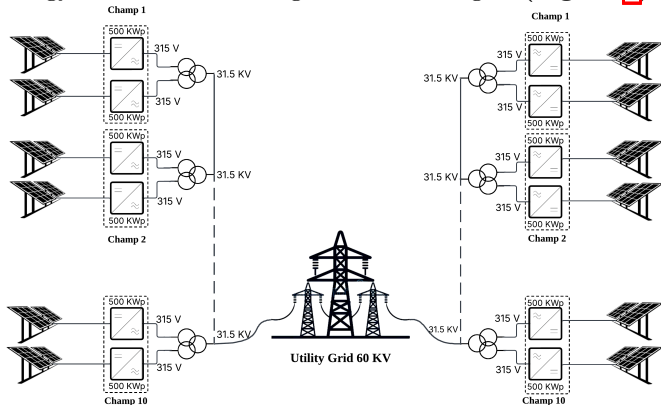


FIGURE 1: OVERVIEW OF THE 20 MW AIN EL-MELH GRID-CONNECTED POWER PLANT

The PV modules are connected to 500 kW inverter cabinets through junction boxes, which serve as the primary source of data collection. The data was gathered over a four-year period, from January 1, 2020, to December 31, 2023, with measurements recorded at 15-minute intervals, resulting in a dataset comprising 69,195 data points. This dataset includes a wide range of parameters such as solar panel temperature, tilt radiation, total radiation, diffuse radiation, direct radiation, wind speed, humidity, pressure, voltage, current, and PV power. An overview of the environmental and electrical parameters of the PV system is presented in Table 1.

TABLE 1: SUMMARY OF ENVIRONMENTAL AND ELECTRICAL PARAMETERS OF THE PV SYSTEM

Features	Description	Min	Max	Average
T_p	Panel temperature ($^{\circ}\text{C}$)	-2.5	74.8	27.9
G_{tilt}	Tilt irradiation (W/m^2)	0	1566	308
G_{tot}	Total irradiation (W/m^2)	0	1396	238
G_{disp}	Dispersion irradiation (W/m^2)	0	648	74
G_{dir}	Direct irradiation (W/m^2)	0	1364	234
WS	Wind speed (m/s)	0	22.2	3.8
H	Humidity (%)	0	71.6	35.2
P	Pressure (Pa)	0	927	911

Features	Description	Min	Max	Average
V_{DC}	Voltage (V)	0	780	323
I_{DC}	Current (A)	0	944	183
P_{DC}	Power (kW)	0	569	108

2.2 Data Processing

The data processing phase involved several critical steps to ensure the quality and usability of the dataset for predictive modeling. First, the raw data, collected from the inverter cabinets and junction boxes, was preprocessed to correct any inconsistencies, handle missing values, and remove outliers. Data points with significant errors or gaps were either interpolated or excluded, ensuring the dataset was both complete and reliable. A key innovation in this study was the correction of the tilt radiation data. During the site measurements, it was observed that the tilt radiation sensor was poorly positioned, causing shading effects that were not present on the photovoltaic modules themselves. To address this, the tilt radiation values were adjusted by scaling them according to the variation in current, as both parameters were recorded simultaneously. This correction ensured that the tilt radiation data more accurately reflected the true solar exposure received by the modules. Additionally, the data underwent normalization to scale the values within a standardized range, which improved the performance of the proposed models by eliminating any biases introduced by differing measurement units or magnitudes. Feature selection was also applied to identify the most relevant variables for the prediction tasks, removing redundant or irrelevant features. Various feature selection techniques, such as correlation analysis and importance ranking, were used to retain only the most influential parameters, such as solar panel temperature, radiation, and environmental conditions, for predicting power, current, and voltage. Following these pre-processing steps, the dataset was split into training and testing subsets, with 80% allocated for model training and the remaining 20% reserved for testing the model generalization capabilities. These pre-processing steps were essential in transforming the raw sensor data into a structured and ready-to-use format for effective machine learning predictions.

3. METHODOLOGY

This section presents the methodology developed in this study to enhance the monitoring, prediction, and fault management of large-scale photovoltaic (PV) systems. The proposed approach is structured around three main components. First, the architecture and training process of a hybrid deep learning model, based on a one-layer Convolution Neural Network (CNN) and Long Short-Term Memory (LSTM) are described for accurate PV power output prediction. Second, a robust fault detection and diagnosis procedure is detailed, leveraging the Corrected Performance Ratio (CPR) and the Anomaly Frequency Index (AFI) to differentiate between transient anomalies and persistent faults. Finally, the deployment strategy is outlined, including the design and implementation of a standalone graphical user interface (GUI) that seamlessly integrates both predictive and diagnostic functionalities. The application processes PV system data at the end of each day to deliver concise performance evaluations and actionable insights. Together, these components form a comprehensive, scalable framework aimed at improving the reliability and operational efficiency of PV systems.

3.1 Models Description

Three primary deep learning approaches have shown significant promise in capturing temporal dependencies and patterns in time series data: the one-dimensional Convolution Neural Network (CNN1D), the Long Short-Term Memory (LSTM) network, and the Transformer model. These methods have become the focus of much of the recent research due to their effectiveness in various sequential prediction tasks. More recently, hybrid models that combine the strengths of these architectures, particularly CNNs and LSTMs, have gained attention, aiming to further enhance prediction accuracy by leveraging both local feature extraction and long-term temporal modeling.

3.1.1 Local Feature Learning with CNN

The convolution neural network (CNN) has demonstrated strong performance in capturing localized patterns across various sequential data domains, such as image processing, sensor streams, and time series analysis. A one-dimensional convolution layer (Conv1D) extends this capability to temporal sequences by applying learnable filters over fixed-size sliding windows, enabling the automatic extraction of key temporal features such as peaks, transitions, or periodic structures—that are critical when sequence order and timing are relevant [21], as illustrated in Figure 2. Mathematically, given an input sequence $\mathbf{x} \in \mathbb{R}^N$ and a convolution filter (kernel) $\mathbf{w} \in \mathbb{R}^K$, the convolution operation produces an output sequence $\mathbf{y} \in \mathbb{R}^{N-K+1}$, where each element is defined as Equation (1) :

$$y[i] = \sum_{j=0}^{K-1} w[j] \cdot x[i+j], \quad \text{for } i = 0, 1, \dots, N-K \quad (1)$$

$$y[j] = (\mathbf{x} * \mathbf{w})[j]$$

where:

- $\mathbf{x} \in \mathbb{R}^N$ is the input sequence of length N ,
- $\mathbf{w} \in \mathbb{R}^K$ is the convolution kernel (filter) of size K ,
- $\mathbf{y} \in \mathbb{R}^{N-K+1}$ is the output feature map of length $N-K+1$.

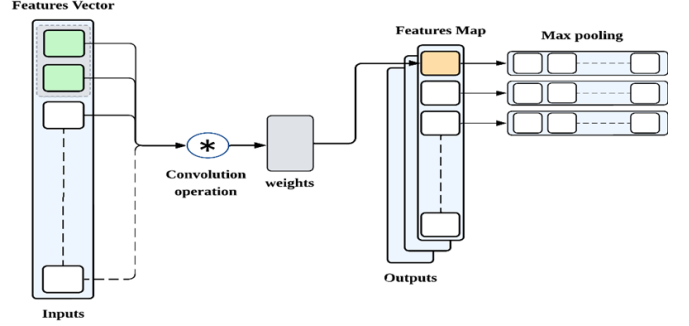


FIGURE 2: ARCHITECTURE OF A CONV1D WITH MAX-POOLING PROCESS

Conv1D has some limitations. It primarily focuses on short-term dependencies and may struggle to capture long-range relationships in the data unless it is complemented by additional layers such as recurrent neural networks (RNNs), long short-term memory networks (LSTMs), or Transformer models. Furthermore, effective use of Conv1D requires careful pre-processing to structure the input data into sequences compatible with convolution operations. To prevent local feature vectors from being used independently, we aggregate them using a special sub-sampling technique called *pooling*. Notably, pooling does not discard any useful information [22]. To control overfitting and reduce computational load, the extracted feature maps are passed through a pooling layer—typically max-pooling—defined by Equation (2), which retains only the most prominent activation within non-overlapping windows of fixed length P . Concretely, if the convolution layer produces a one-dimensional feature map:

$$\mathbf{y} = [y_0, y_1, \dots, y_{N-K}] \in \mathbb{R}^{N-K+1}$$

Then, max pooling over windows of size P yields a pooled feature vector:

$$\mathbf{p} = [p_0, p_1, \dots, p_{L-1}] \in \mathbb{R}^L, \quad \text{where } L = \lfloor \frac{N-K+1}{P} \rfloor$$

Where: and each element of is given by:

$$p_l = \max(y_{lP}, y_{lP+1}, \dots, y_{lP+P-1}), \quad \text{for } l = 0, 1, \dots, L-1 \quad (2)$$

In other words, for each pooling window $\{y_{lP+i}\}_{i=0}^{P-1}$, max pooling selects the single largest activation. This operation effectively maps:

$$\mathbb{R}^{N-K+1} \longrightarrow \mathbb{R}^{\lfloor \frac{N-K+1}{P} \rfloor}$$

thus reducing dimensionality while preserving the most salient features in each local region of the sequence.

3.1.2 LSTM Layer

Long Short-Term Memory (LSTMs) networks, introduced by Hochreiter and Schmidhuber, are an advanced form of Recurrent Neural Networks (RNNs) specifically designed to cap-

ture long-term dependencies in sequential data. Unlike traditional RNNs, which pass a single hidden state through time and often struggle with vanishing gradients, LSTMs incorporate a dedicated memory cell that enables them to retain and manage information over extended sequences [23]. This architectural enhancement (see Figure 3) allows LSTMs to effectively learn temporal patterns and relationships that span long intervals, making them highly suitable for tasks such as language translation, speech recognition, and time series forecasting.

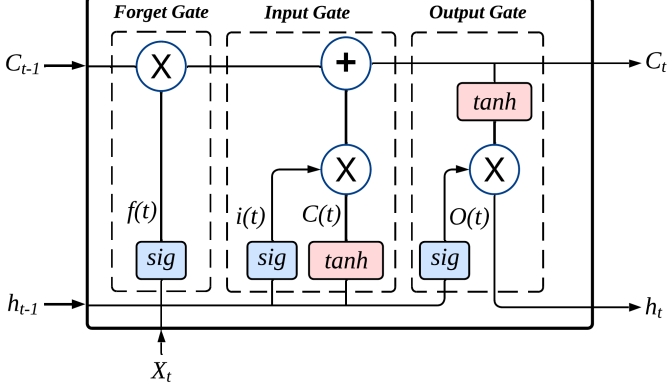


FIGURE 3: THE LSTM ARCHITECTURE

An LSTM network is built around a central memory cell that maintains information over time. This memory is regulated by three primary gates:

- **Forget Gate:** Decides which information should be discarded from the cell state. It is computed as:

$$\mathbf{f}_t = \sigma(\mathbf{W}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) \quad (3)$$

- **Input Gate:** Determines which new information is added to the cell state. It is computed in two parts:

$$\mathbf{i}_t = \sigma(\mathbf{W}_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i) \quad (4)$$

$$\tilde{\mathbf{C}}_t = \tanh(\mathbf{W}_c \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c) \quad (5)$$

- **Output Gate:** Controls what information is output from the memory cell:

$$\mathbf{o}_t = \sigma(\mathbf{W}_o \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o) \quad (6)$$

Through these gates, LSTMs can selectively retain, forget, or output information, enabling powerful sequence learning capabilities over long durations.

3.1.3 CNN-LSTM Hybrid Model

The proposed CNN-LSTM model, shown in Figure 4, includes 1D-CNN layer, one LSTM layer, and a fully connected layer, combining feature extraction and temporal learning to improve prediction accuracy. While a standard RNN can process sequential data, it often forgets long-term dependencies. The LSTM component overcomes this limitation by preserving both short and long-term memory.

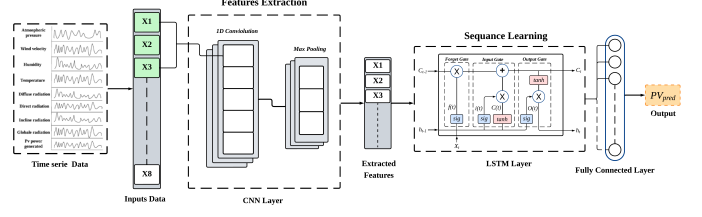


FIGURE 4: THE PROPOSED CNN-LSTM ARCHITECTURE

The input of the neural network consists of ten variables, including temperature, irradiance, global radiation, and others (see the Dataset section for more details). The upper layers of the CNN-LSTM model begin with a convolution component, which extracts local spatial patterns from variables that influence PV power generation, such as temperature, radiation intensity, and ambient conditions.

The dataset is divided into two parts: 80% is used for training the model, and 20% is reserved for validation. The CNN architecture includes an input layer that receives sensor data, several hidden layers (including a 1-D convolution layer and a max-pooling layer), and an output layer that transfers feature representations to the LSTM. While the CNN focuses on capturing local features from input data, the LSTM specializes in learning temporal dependencies across time steps.

Through this architecture, the model learns an optimized mapping between input variables and predicted outputs via weight adjustment. Model hyperparameters such as the number of layers, filter sizes, and learning rates were tuned empirically using trial-and-error. For each set of parameters, the model was trained on 80% of the data and validated on the remaining 20%, with performance evaluated seasonally across different environmental conditions.

To assess the performance of the proposed CNN-LSTM model, we compare it with standalone LSTM, a deeper CNN-LSTM variant (5D-CNN-LSTM), and a Random Forest (RF) model [24]. The results of these comparisons are detailed in later sections.

Model accuracy was quantified using three standard regression metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). These metrics are formally defined in Equations (7)–(9), and offer insight into the predictive precision and overall goodness-of-fit of each model [25, 26, 27]:

$$\text{MAE} = \frac{1}{n} \sum_{i=0}^{n-1} |y_i - \hat{y}_i| \quad (7)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2} \quad (8)$$

$$R^2 = 1 - \frac{\sum_{i=0}^{n-1} (\hat{y}_i - y_i)^2}{\sum_{i=0}^{n-1} (\bar{y} - y_i)^2}, \quad \text{where } \bar{y} = \frac{1}{n} \sum_{i=0}^{n-1} y_i \quad (9)$$

The evaluation results were then analyzed and compared to identify the best-performing model for PV generation prediction. This analysis highlighted the strengths and weaknesses of each approach and guided the selection of the most suitable model.

3.2 Fault Detection and Diagnosis Method

Photovoltaic (PV) systems are subject to a range of faults that can compromise energy yield and operational reliability. To ensure continuous and efficient performance, early identification and accurate classification of faults are essential. This section presents an integrated two-stage framework that first detects anomalies in system behavior and then diagnoses their underlying causes. The detection phase identifies deviations in inverter performance based on a statistical analysis of corrected performance metrics, while the diagnosis phase classifies fault types using adaptive predictive boundaries for current and voltage derived from a data-driven CNN-LSTM prediction model. The entire process of the proposed method is outlined in the flowchart shown in Figure 5.

3.2.1 Fault Detection Approach

The proposed fault detection approach is designed to efficiently identify anomalies in large-scale photovoltaic (PV) systems by analyzing the Corrected Performance Ratio (CPR) in conjunction with statistical outlier detection and temporal fault tracking. This method integrates three key components: CPR analysis, the Hampel identifier, and the Anomaly Frequency Index (AFI), forming a comprehensive framework capable of distinguishing between transient anomalies and persistent system faults. The detection process unfolds in the following structured steps: The proposed fault detection approach is designed to efficiently identify anomalies in large-scale photovoltaic (PV) systems by analyzing the Corrected Performance Ratio (CPR) in conjunction with statistical outlier detection and temporal fault tracking. This method integrates three key components: CPR analysis, the Hampel identifier, and the Anomaly Frequency Index (AFI), forming a comprehensive framework capable of distinguishing between transient anomalies and persistent system faults. The detection process unfolds in the following structured steps:

A. Performance Metric

The IEC 61724-1 standard defines the Performance Ratio (PR) as a key metric for evaluating the efficiency of solar power plants. A drop in the PR value typically indicates a fault condition, enabling operators to identify deviations from normal system performance. PR is defined as the ratio between the final energy yield and the reference yield:

$$PR = \frac{Y_f}{Y_r} \quad (10)$$

where Y_f is the final yield and Y_r is the reference yield.

To account for weather-related variations in PV system performance, the National Renewable Energy Laboratory (NREL) [28] developed the Corrected Performance Ratio (CPR). It is calculated as:

$$CPR = \frac{PR}{1 - \delta(T_{\text{cell_Type_avg}} - T_{\text{cell}})} \quad (11)$$

Where:

- T_{cell} : Cell temperature computed from measured meteorological data ($^{\circ}\text{C}$)

- $T_{\text{cell_Type_avg}}$: Average cell temperature computed from one year of weather data using the project weather file ($^{\circ}\text{C}$)
- δ : Temperature coefficient for power (

This correction ensures that CPR reflects true performance degradation, independent of natural fluctuations in solar irradiance or temperature.

B. Abnormal Behavior Detection – Hampel Identifier

To detect faults in PV inverters, the Hampel identifier is applied as a robust statistical tool for outlier detection. It identifies abnormal behavior by flagging data points that fall outside a defined threshold range. Specifically, an inverter's CPR value at time t , denoted as CPR_t^i for inverter i , is considered an outlier if it drops below the lower boundary established by the Hampel identifier:

$$CPR_t^i < \text{bound} \quad (12)$$

where the bound is calculated as:

$$\text{bound} = \mu_{CPR_t} - \alpha \cdot S \quad (13)$$

$$S = \frac{1}{0.6745} \cdot \text{median} \left| \mu_{CPR_t} - CPR_t^i \right| \quad (14)$$

with α being a threshold constant (typically 3), and S the median absolute deviation (MAD). To ensure numerical stability during periods of low irradiance (e.g., sunrise, sunset), the MAD is adjusted as $\max(S, 0.07)$.

C. Fault Detection Frequency – Anomaly Frequency Index (AFI)

While individual inverter CPR outliers may indicate faults, short-term anomalies can also result from transient environmental fluctuations or measurement noise. To differentiate between isolated events and persistent faults, the Anomaly Frequency Index (AFI) is introduced:

$$AFI_i = \frac{1}{M} \sum_{t=1}^M H(\text{bound}_t - CPR_t^i) \quad (15)$$

Here, bound_t is the Hampel lower bound at time t , and $H(x)$ is the Heaviside step function:

$$H(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (16)$$

Each 15-minute interval is flagged as 1 (abnormal) or 0 (normal), and AFI is computed over non-overlapping 75-minute windows (5 intervals). If $AFI \geq \tau$ (e.g., $\tau = 0.6$), the window is labeled as “faulty”; otherwise, it is “healthy.” This approach enables temporal fault tracking and reduces false alarms.

3.2.2 Fault Diagnosis Approach

Once the fault detection approach identifies anomalous behavior (via the Anomaly Fault Indicator, AFI), indicating a potential fault, the next step is to diagnose its root cause. To isolate and determine the fault type, we propose a novel method based on adaptive current and voltage boundaries, as follows: The trained models predict the expected DC current and voltage ($I_{dc,pred}$ and $V_{dc,pred}$) under fault-free, fully operational conditions for the given day. Adaptive boundaries are then established based on these predictions using the following equations.

Once fault detection identifies anomalous behavior (via the Anomaly Fault Indicator, AFI), indicating a potential fault, the next step is diagnosing its root cause. To isolate and determine the fault type, we propose a novel method based on adaptive current and voltage boundaries, as follows: The trained models predict the expected DC current and voltage ($I_{DC,pred}$ and $V_{DC,pred}$) under fault-free, fully operational conditions for the given day. Adaptive boundaries are then established based on these predictions using the following equations:

$$I_{DC,lower} < I_{DC,measured} < I_{DC,upper} \quad (17)$$

$$V_{DC,lower} < V_{DC,measured} < V_{DC,upper} \quad (18)$$

With:

$$I_{DC,lower} = I_{DC,pred} - p \cdot \sigma_I \quad (19)$$

$$I_{DC,upper} = I_{DC,pred} + p \cdot \sigma_I \quad (20)$$

$$V_{DC,lower} = V_{DC,pred} - p \cdot \sigma_V \quad (21)$$

$$V_{DC,upper} = V_{DC,pred} + p \cdot \sigma_V \quad (22)$$

where σ_I and σ_V are the standard deviations of prediction errors (i.e., $|\text{measured} - \text{predicted}|$) over the day, and p is a tuning factor.

During real-time operation, if $I_{DC,measured}$ or $V_{DC,measured}$ fall outside these adaptive bounds, a fault is flagged. This enables detection of current drops or voltage dips due to string or module failures. The system classifies the fault source by evaluating whether current and voltage measurements breach the respective adaptive boundaries.

The following scenarios help identify possible PV system fault types:

- **Scenario 1:** $I_{DC,measured}$ and $V_{DC,measured}$ both exceed their threshold bands. This deviation strongly suggests partial shading or soiling effects on the PV array, causing a simultaneous drop in current and voltage due to reduced irradiance and increased mismatch losses.
- **Scenario 2:** $I_{DC,measured}$ remains within the threshold band, but $V_{DC,measured}$ is outside. This may indicate a short-circuit fault in the PV modules.
- **Scenario 3:** $I_{DC,measured}$ is outside the threshold band, while $V_{DC,measured}$ remains within. This suggests a disconnection issue affecting PV modules or strings.

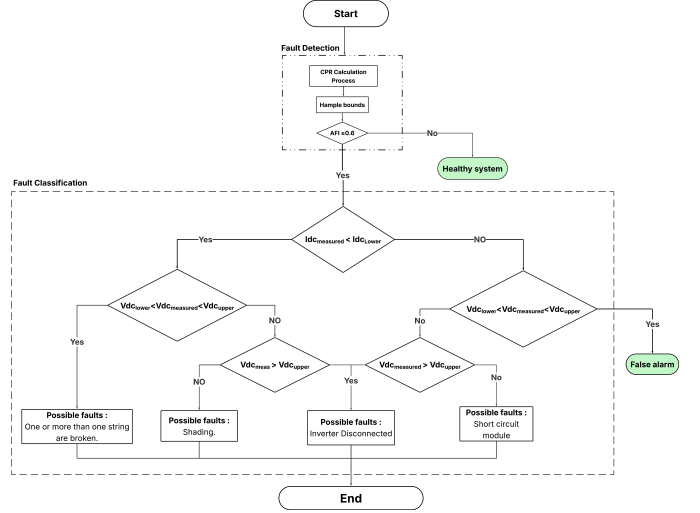


FIGURE 5: FAULT DETECTION AND DIAGNOSIS FLOWCHART OF THE PROPOSED METHOD

3.3 Standalone GUI Application

The final stage of this work centers on deploying the best-performing CNN-LSTM (Convolution Neural Network–Long Short-Term Memory) model within a standalone Python-based GUI application. This application tool operates locally on individual machines without the need for external dependencies or internet access, delivering reliable prediction of PV system output.

In addition to prediction capabilities, the backend incorporates computation modules designed to assess the system's performance. These modules calculate normalized energy yields (energy output per installed kilowatt), quantify key system losses, and determine important performance indices such as the Performance Ratio (PR) and the Corrected Performance Ratio (CPR), accounting for temperature effects on module efficiency. The CPR metric also plays a central role in the fault detection process.

By processing data at the end of the day, these analytics provide immediate insights into the system's operational status, flag anomalies as they occur, and display alerts in the interface application. By combining the predictive power of the CNN-LSTM model with standardized performance analysis, the platform delivers a comprehensive overview of plant health. This seamless integration enhances real-time situational awareness and adds significant value by aligning predictive forecasts with actual operational metrics, thereby supporting more informed monitoring and decision-making.

4. RESULTS AND DISCUSSION

4.1 Prediction Results

This section presents the results obtained from the datasets, emphasizing the predicted performance of the PV system under various weather conditions. The prediction values for power, current, and voltage are evaluated against real-world measurements. The evaluation is made considering a dataset collected during one operating day. Additionally, we used a dataset that spanned 12 days, one from each month of the year, to assess the

prediction model's ability to capture the long-term variability introduced by different seasons and environmental conditions.

4.1.1 DC Power Prediction

Figures 7, 8, and 9 illustrate the predicted and actual output current for the three test scenarios: a cloudy day, a clear-sky day, and a representative 12-day period covering seasonal variability. The results demonstrate that the CNN-LSTM model achieves high accuracy in forecasting daily PV output, with R^2 values consistently exceeding 0.9992 across different weather conditions, as detailed in Table 2.

The Mean Absolute Error (MAE) decreased from 2.4475 kW under cloudy sky conditions to 1.3795 kW under clear skies. Similarly, the Root Mean Square Error (RMSE) dropped from 3.89 kW in cloudy conditions to 2.17 kW on clear days. This improvement in predictive accuracy is primarily attributed to the reduced fluctuation in power output during clear-sky conditions, where irradiance and temperature levels are more stable throughout the day.

Overall, the daily performance indicators validate the model's reliability, especially considering that the PV system's peak output exceeds 400 kW.

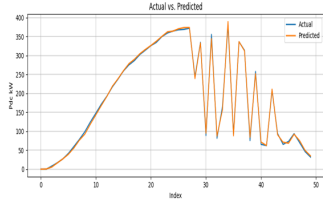


FIGURE 7: PREDICTED AND MEASURED POWER CLOUDY DAY

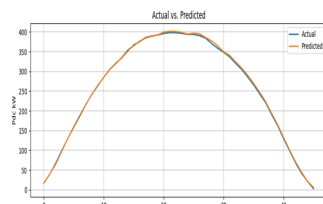


FIGURE 8: PREDICTED AND MEASURED POWER CLEAR DAY

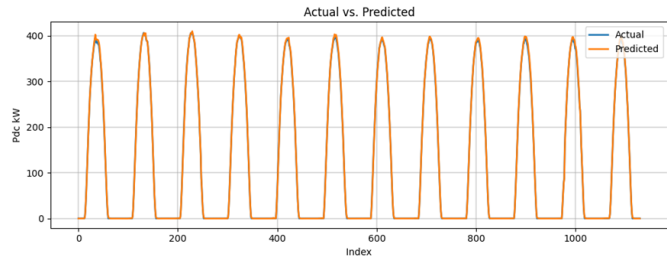


FIGURE 9: PREDICTED AND MEASURED POWER FOR 12-DAY PERIOD OF THE YEAR.

TABLE 2: REGRESSION METRICS OF PV POWER PREDICTION USING CNN-LSTM MODEL.

Condition	R2	MSE	MAE
Clear day	0.9998	2.1774	1.3795
cloudy day	0.9992	3.8901	2.4475
12-day period	0.9998	2.0872	1.1995

The prediction performance of the output power over a 12-day period demonstrates highly consistent results, with an R^2 value of 0.9998 and an MAE of 1.1995 kW. These findings indicate that the model delivers accurate predictions and is effectively adapted to capturing long-term seasonal variability.

4.1.2 DC Current prediction

The three test scenarios were also conducted to evaluate the proposed model's performance in predicting the DC current generated by the PV system. The results, shown in Figures 11, 10, and 12 demonstrate that the CNN-LSTM model predictions closely match the actual measurements. The model achieved R^2 values exceeding 0.9991 across both daily scenarios, as summarized in Table 3.

The Mean Absolute Error (MAE) ranged from 2.2967 A to 5.1988 A, while the Root Mean Square Error (RMSE) fell within the interval of 3.5286 A to 8.2184 A. These results confirm the model's robustness and accuracy in current prediction, especially considering that the DC current can reach values up to 800 A during peak solar production.

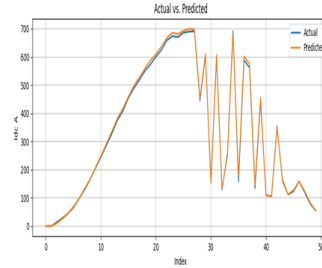


FIGURE 10: PREDICTED AND MEASURED CURRENT CLOUDY DAY

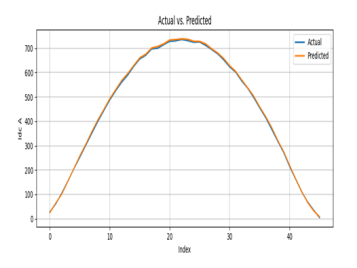


FIGURE 11: PREDICTED AND MEASURED CURRENT CLEAR DAY

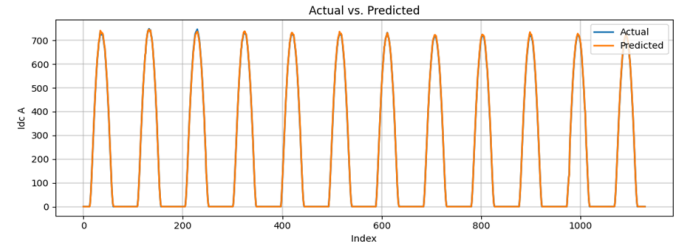


FIGURE 12: PREDICTED AND MEASURED CURRENT 12-DAY PERIOD OF THE YEAR

TABLE 3: REGRESSION METRICS OF PV CURRENT PREDICTION USING CNN-LSTM MODEL.

Condition	R2	MSE	MAE
Clear day	0.9998	3.5286	2.2967
cloudy day	0.9991	8.2184	5.1988
12-day period	0.9999	3.3540	2.0399

On the other hand, the long-term predicted results obtained over a 12-day period show practically similar performance compared to daily predictions in terms of R^2 evaluated metrics.

4.1.3 DC Voltage prediction

Figures 13, 14 present the predicted DC voltage generated by the CNN-LSTM model alongside the measured values under both clear and cloudy sky conditions. As shown, the predicted voltage closely aligns with the actual measurements. This consistency is further supported by the regression results summarized in Table 4, where the coefficient of determination (R^2) reaches approximately 0.9660, indicating a strong model fit.

These results validate the robustness and adaptability of the proposed model, particularly under stable environmental conditions.

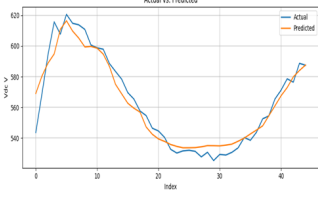


FIGURE 13: PREDICTED AND MEASURED VOLTAGE CLOUDY DAY

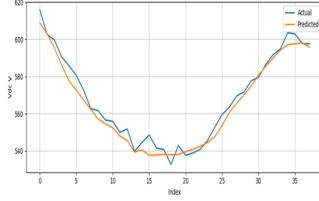


FIGURE 14: PREDICTED AND MEASURED VOLTAGE CLEAR DAY

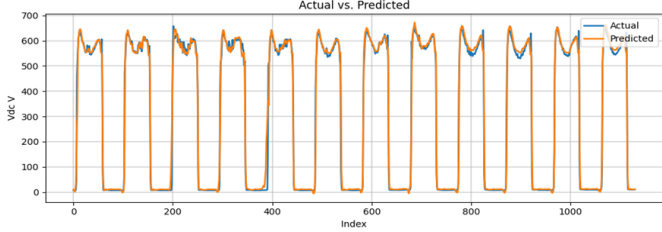


FIGURE 15: PREDICTED AND MEASURED VOLTAGE 12-DAY PERIOD OF THE YEAR

TABLE 4: REGRESSION METRICS OF PV VOLTAGE PREDICTION USING CNN-LSTM MODEL.

Condition	R2	MSE	MAE
Clear day	0.9660	4.3260	3.5461
cloudy day	0.9414	7.1951	5.3641
12-day period	0.9869	13.2100	1.1995

The 12-day long-term period predictions demonstrate practically the same R^2 performance as the daily predictions, Figures 15. To justify our selection of a single-layer CNN-LSTM model, we compared its performance against several alternatives: a standalone LSTM, a five-layer CNN-LSTM, and a Random Forest. Because our primary goal is PV system fault diagnosis, we restrict our analysis to daily power predictions, which are most relevant for short-term diagnostic tasks. Table 5 summarizes the key performance metrics R^2 , RMSE, and MAE obtained from each of these machine learning techniques.

TABLE 5: COMPARATIVE PERFORMANCE ANALYSIS OF PV POWER GENERATION.

Model	R2	MSE	MAE	N parameters
CNN-LSTM	0.9998	3.5286	2.2967	169,793
LSTM	0.9957	11.9234	7.9842	51,821
5CNN-LSTM	0.9975	8.9992	5.4994	1,089,713
RF	0.968	21.02	7.40	None

As shown in Table 5, the single-layer CNN-LSTM model clearly outperforms the other evaluated models by effectively combining spatial feature extraction through CNN with temporal sequence learning via LSTM. In contrast, increasing the number of convolution layers (as in the 5-CNN-LSTM model) leads to a substantial rise in the number of trainable parameters, which may reduce real-time inference efficiency, increase computational requirements, and introduce potential training instability. Additionally, the Random Forest (RF) model ex-

hibits noticeably lower predictive performance, consistent with the findings reported in 24.

4.2 Faults detection and diagnosis results

The experimental validation of the fault detection and diagnosis methods was carried out on a 500 kWp grid-connected photovoltaic system situated in Ain El Melh, Algeria. Throughout this process, several fault types were accurately detected, including a disconnected string, a short-circuited module, and partial shading.

4.2.1 Healthy PV system day

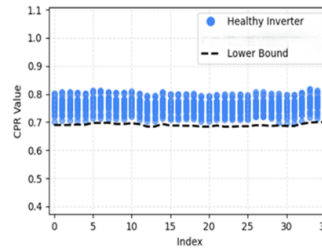


FIGURE 16: FAULT DETECTION VIA CPR ANALYSIS

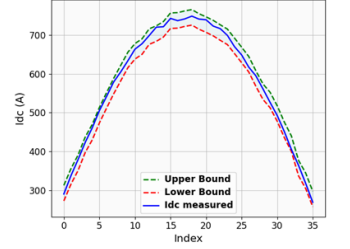


FIGURE 17: FAULT DIAGNOSIS VIA ADAPTIVE CURRENT BOUNDARIES

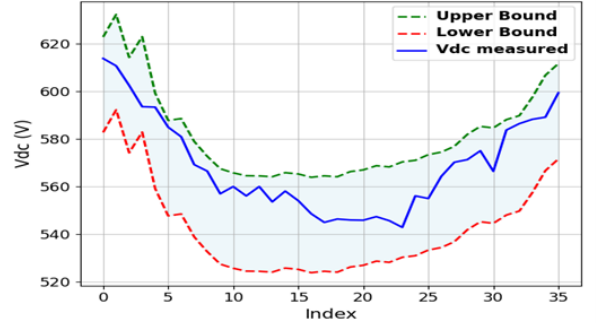


FIGURE 18: FAULT DIAGNOSIS VIA ADAPTIVE VOLTAGE BOUNDARIES

Figure 16 illustrates the variation of CPR values across 40 inverters throughout a single day. The black dashed lines represent the dynamic lower threshold established by the Hampel identifier at each time step. As observed, the majority of inverters exhibit healthy behavior, with CPR values consistently remaining above this threshold.

In cases where red points appear, indicating brief dips of CPR values below the Hampel boundary, these are not considered critical if the corresponding Anomaly Frequency Index (AFI) values remain below the defined threshold for persistent faults. Such instances are interpreted as transient anomalies, likely resulting from short-term factors such as partial shading.

In addition, the adaptive current and voltage boundaries, shown in Figures 17 and 18, remain within their expected limits, reflecting stable electrical performance. This overall consistency confirms the absence of active faults, and no alarms are triggered during the observed period. These findings provide a reliable baseline for future anomaly detection and strengthen the robustness of the fault diagnosis framework for photovoltaic

systems.

4.2.2 Strings disconnected in one inverter

The disconnection of strings from one inverter in the PV array exhibits faulty behavior, with CPR values consistently falling below the Hampel threshold for most of the observed period, as illustrated in Figure 19.

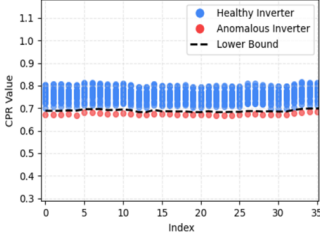


FIGURE 19: FAULT DETECTION VIA CPR ANALYSIS

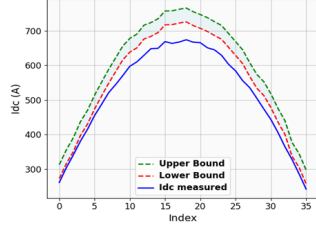


FIGURE 20: FAULT DIAGNOSIS VIA ADAPTIVE CURRENT BOUNDARIES

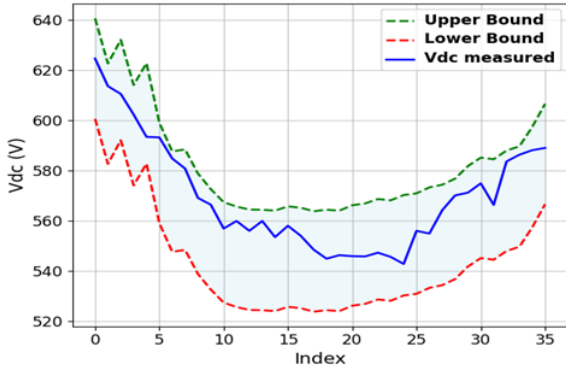


FIGURE 21: FAULT DIAGNOSIS VIA ADAPTIVE VOLTAGE BOUNDARIES

Under this fault condition, the measured current deviates beyond its predefined lower boundary, as shown in Figure 20 while the measured voltage remains within permissible limits (Figure 21). This distinctive pattern—characterized by a significant current deviation accompanied by stable voltage behavior—helps distinguish the fault from other system anomalies. As a result, the diagnosis process accurately identifies the issue as a string disconnection.

4.2.3 Short circuited modules

Short-circuited modules in one inverter string exhibit abnormal behavior, with CPR values consistently falling below the Hampel threshold, as shown in Figure 22 indicating significant power loss or potential system failure. In parallel, two distinct diagnostic indicators are observed: the measured current remains largely within acceptable bounds, while the voltage exhibits a complete deviation beyond its operational limits (Figures 23 and 24). This combination of patterns provides strong diagnostic evidence, allowing for the precise classification of the fault as short-circuited modules within the PV array.

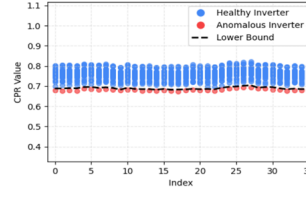


FIGURE 22: FAULT DETECTION VIA CPR ANALYSIS

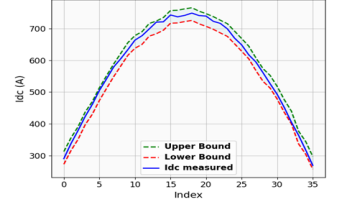


FIGURE 23: FAULT DIAGNOSIS VIA ADAPTIVE CURRENT BOUNDARIES

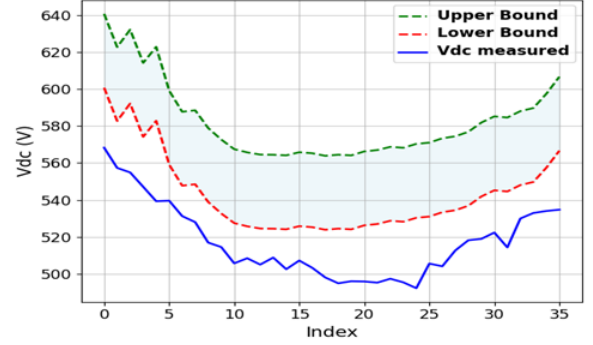


FIGURE 24: FAULT DIAGNOSIS VIA ADAPTIVE VOLTAGE BOUNDARIES

4.2.4 Shading faults

Figure 25 displays abnormal CPR behavior under partial shading conditions, with values consistently falling below the Hampel threshold. In this scenario, both the measured current and voltage exceed their respective adaptive boundaries, as shown in Figures 26 and 27, triggering fault alert signals. The convergence of these indicators allows the anomaly to be accurately classified as a shading-induced fault within the PV system.

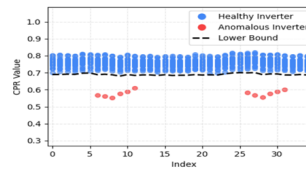


FIGURE 25: FAULT DETECTION VIA CPR ANALYSIS

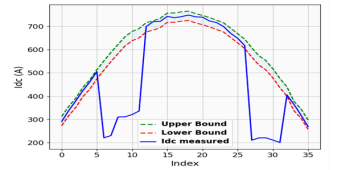


FIGURE 26: FAULT DIAGNOSIS VIA ADAPTIVE CURRENT BOUNDARIES

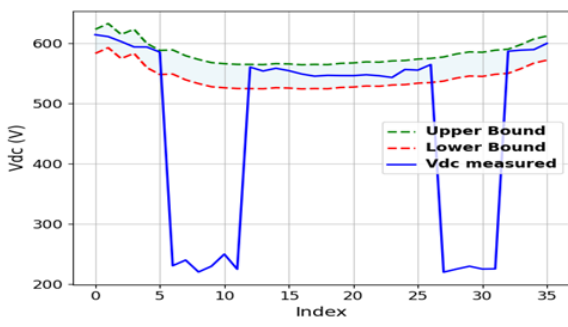


FIGURE 27: FAULT DIAGNOSIS VIA ADAPTIVE VOLTAGE BOUNDARIES

4.3 Standalone Python-based GUI application

This section presents the design and implementation of an intuitive standalone GUI application for power prediction, fault detection, and diagnosis. The application features a user-friendly interface that seamlessly integrates with advanced algorithms, enabling users to:

- Input relevant system data
- Configure prediction parameters
- Visualize immediate measurements alongside predictions
- Perform parameter calculation
- Detect and diagnose system faults

Key features of the developed application include the following:

- **Tab-Based Navigation:** Organized interface with dedicated tabs for prediction tasks, parameter calculations, and fault detection/diagnosis.
- **Interactive Controls:** Intuitive buttons, toggle switches, and state buttons for task initiation and parameter customization.
- **Advanced Visualization:** Dynamic graphical displays (including UI Axes components) for clear representation of system data and analysis results.
- **Data Export Capability:** Option to export results for further analysis or system integration.
- **Seamless Data Integration:** Direct processing of PV station Excel files for real-time predictions without pre-processing.

The application provides specialized tabs for various functions, including PV power generation prediction, yield and loss calculations, and comprehensive fault detection and diagnosis, as illustrated in Figures 28, 29, and 30. This streamlined design ensures efficient monitoring and analysis while maintaining operational simplicity.

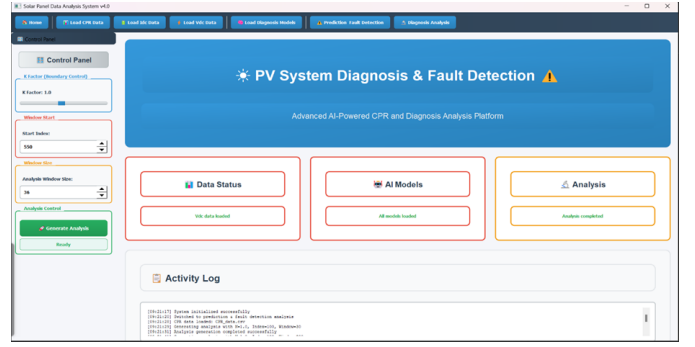


FIGURE 28: MAIN STANDALONE GUI INTERFACE MONITORING



FIGURE 29: PREDICTION AND FAULT DETECTION MODULE

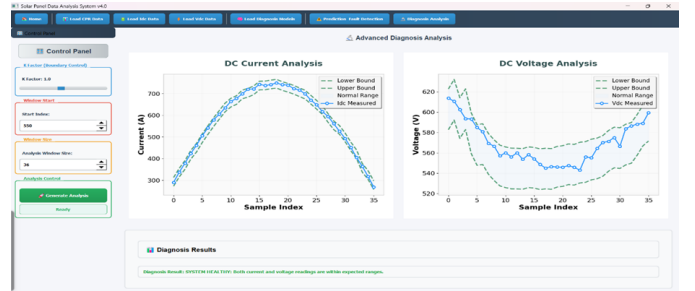


FIGURE 30: FAULT DIAGNOSIS MODULE

5. CONCLUSION

This study presents a comprehensive procedure for automated supervision and fault diagnosis in grid-connected photovoltaic (PV) systems. It integrates a complete monitoring and simulation package within a unified environment, demonstrating the effectiveness of deep learning techniques in accurately predicting PV power generation across the entire operating system. Through careful data pre-processing, feature selection, and model evaluation, CNN- LSTM models were identified as the most effective for estimating power, current, and voltage outputs from PV systems installed in Algeria. For fault detection, reference thresholds were established using the Corrected Performance Ratio (CPR) method alongside the Anomaly Frequency Index (AFI), enabling the identification of inverters that exhibit frequent anomalies over time. For fault diagnosis, adaptive boundaries for current and voltage

deviations were defined to determine whether faults origi-

nate from current reductions or voltage drops, facilitating the identification of the most probable fault causes. Moreover, beyond fault detection and diagnosis, the study proposes the integration of the best-performing predictive model, along with computation modules for assessing energy yield, system losses, and corrected performance ratios, into a Standalone Python-based GUI application to enable immediate predictions and monitoring. This integrated analysis framework provides actionable insights that enhance performance optimization and enable early detection and diagnosis of potential faults in photovoltaic systems.

ACKNOWLEDGMENT

The author would like to express heartfelt gratitude to Professor Aissa Chouder for his generous guidance, support, and invaluable feedback throughout this work. Special thanks are also extended to Professor Houcine Oudira for his encouragement and supervision.

ETHICAL STATEMENT

This study was conducted in accordance with the principles of scientific integrity .

REFERENCES

- [1] GlobalData, *Solar Photovoltaic (PV) Market, Update 2019—Global Market Size, Market Share, Average Price, Regulations, and Key Country Analysis to 2030*, London, UK: GlobalData, 2019.
- [2] Y. Wang *et al.*, “Accelerating the energy transition towards photovoltaic and wind in China,” *Nature*, vol. 619, no. 7971, pp. 761–767, 2023, doi: 10.1038/s41586-023-06180-8.
- [3] A. Jaeger-Waldau, *PV Status Report 2019, EUR 29938 EN*, Luxembourg: Publications Office of the European Union, 2019.
- [4] IEA, “Renewables 2021 Analysis and Forecast to 2026,” Int. Energy Agency, 2021. [Online]. Available: <https://webstore.iea.org/download/direct/4329>
- [5] Y. Gaaloul *et al.*, “Faults Detection and Diagnosis of a Large-Scale PV System by Analyzing Power Losses and Electric Indicators Computed Using Random Forest and KNN-Based Prediction Models,” *Energies*, vol. 18, p. 2482, 2025. doi: 10.3390/en18102482.
- [6] A. F. Amiri, H. Oudira, A. Chouder, and S. Kichou, “Faults detection and diagnosis of PV systems based on machine learning approach using random forest classifier,” *Energy Convers. Manag.*, vol. 301, p. 118076, 2024, doi: 10.1016/j.enconman.2024.118076.
- [7] Y. Ledmaoui *et al.*, “Enhancing Solar Power Efficiency: Smart Metering and ANN-Based Production Forecasting,” *Computers*, vol. 13, no. 9, pp. 1–18, 2024, doi: 10.3390/computers13090235.
- [8] W. Shinong, M. Qianlong, X. Jie, G. Yuan, and L. Shilin, “An improved mathematical model of photovoltaic cells based on datasheet information,” *Sol. Energy*, vol. 199, pp. 437–446, 2020, doi: 10.1016/j.solener.2020.02.046.
- [9] M. J. Mayer and G. Gróf, “Extensive comparison of physical models for photovoltaic power forecasting,” *Appl. Energy*, vol. 283, p. 116239, 2021.
- [10] R. Haji, M. Zargar, M. H. Yaghmaee, and S. Member, “Development of a Markov-Chain-Based Solar Generation Model for Smart Microgrid Energy Management System,” *IEEE Trans. Sustain. Energy*, vol. 11, pp. 736–745, 2020.
- [11] M. Gao, K. Wang, and A. C. T. M. Chain, “Probabilistic Model Checking for Green Energy Router System in Energy Internet,” in *Proc. GLOBECOM 2017*, Singapore, Dec. 2017.
- [12] J. Wang *et al.*, “A Short-Term Photovoltaic Power Prediction Model Based on the Gradient Boost Decision Tree,” *Appl. Sci.*, vol. 8, p. 689, 2018.
- [13] A. Dong *et al.*, “Fault diagnosis and classification in photovoltaic systems using SCADA data,” in *Proc. Int. Conf. Sensing, Diagnostics, Prognostics, and Control (SDPC)*, Shanghai, China, 2017, pp. 1–6.
- [14] E. Ogliari, A. Dolara, G. Manzolini, and S. Leva, “Physical and hybrid methods comparison for the day ahead PV output power forecast,” *Renew. Energy*, vol. 113, pp. 11–21, 2017, doi: 10.1016/j.renene.2017.05.063.
- [15] A. Chouder and S. Silvestre, “Automatic supervision and fault detection of PV systems based on power losses analysis,” *Energy Convers. Manag.*, vol. 51, no. 10, pp. 1929–1937, 2010, doi: 10.1016/j.enconman.2010.02.025.
- [16] B. Taghezouit *et al.*, “Intelligent Monitoring of Photovoltaic Systems via Simplicial Empirical Models and Performance Loss Rate Evaluation under LabVIEW: A Case Study,” *Energies*, vol. 15, no. 21, 2022, doi: 10.3390/en15217955.
- [17] K. Muralitharan, R. Sakthivel, and R. Vishnuvarthan, “Neural Network based Optimization Approach for Energy Demand Prediction in Smart Grid,” *Neurocomputing*, 2017, doi: 10.1016/j.neucom.2017.08.017.
- [18] Y. Liu *et al.*, “Fault Diagnosis Approach for Photovoltaic Array Based on the Stacked Auto-Encoder and Clustering with I-V Curves,” *Energy Convers. Manag.*, vol. 245, p. 114603, 2021.
- [19] Z. Chen *et al.*, “Deep Residual Network Based Fault Detection and Diagnosis of Photovoltaic Arrays Using Current-Voltage Curves and Ambient Conditions,” *Energy Convers. Manag.*, vol. 198, p. 111793, 2019.
- [20] Y. Ledmaoui *et al.*, “Enhanced Fault Detection in Photovoltaic Panels Using CNN-Based Classification with PyQt5 Implementation,” *Sensors*, vol. 24, no. 22, pp. 1–20, 2024, doi: 10.3390/s24227407.
- [21] S. Kiranyaz *et al.*, “1D convolutional neural networks and applications: A survey,” *Mech. Syst. Signal Process.*, vol. 151, p. 107398, 2021, doi: 10.1016/j.ymssp.2020.107398.
- [22] Y. L. Boureau, J. Ponce, and Y. LeCun, “A theoretical analysis of feature pooling in visual recognition,” in *Proc. ICML*, Haifa, Israel, Jun. 2010, pp. 111–118.
- [23] M. Tovar, M. Robles, and F. Rashid, “PV Power Prediction, Using CNN LSTM Hybrid Neural Network Model: Case Study: Temixco Morelos, México,” *Energies*, vol. 13, no. 24, Art. no. 6578, Dec. 2020.
- [24] A. F. Amiri *et al.*, “Improving Photovoltaic Power Prediction: Insights through Computational Modeling and Feature Selection,” *Energies*, vol. 17, p. 3078, 2024. doi: 10.3390/en17133078.

-
- [25] I. Shah, H. Iftikhar, and S. Ali, “Modeling and Forecasting Electricity Demand and Prices: A Comparison of Alternative Approaches,” *J. Math.*, vol. 2022, Art. no. 3581037, 2022.
- [26] I. Shah, F. Jan, and S. Ali, “Functional data approach for short-term electricity demand forecasting,” *Math. Probl. Eng.*, vol. 2022, Art. no. 6709779, 2022.
- [27] F. Lisi and I. Shah, “Forecasting next-day electricity demand and prices based on functional models,” *Energy Syst.*, vol. 11, pp. 947–979, 2020.
- [28] T. Dierauf *et al.*, “Weather-Corrected Performance Ratio,” NREL, Golden, CO, USA, 2013.