İSTANBUL UNIVERSITY CERRAHPAŞA

REVIEW
Electrica 2023; 23(1): 137-148

A Comparative Analysis of Solar Photovoltaic Advanced Fault Detection and Monitoring Techniques

Siti Nor Azlina Mohd Ghazali, Muhamad Zahim Sujod

Department of Electrical Engineering, College of Engineering, Universiti Malaysia Pahang (UMP), Pekan, Pahang, Malaysia

Cite this article as: S. N. A. M. Ghazali and M. Z. Sujod, "A comparative analysis of solar photovoltaic advanced fault detection and monitoring techniques," *Electrica*, 23(1), 137-148, 2023.

ABSTRACT

The non-linear I-V characteristics of the photovoltaic output have affected fault detection methods to work accurately. This scenario can cause hidden faults in the system and reduces overall productivity. Fault detection and monitoring techniques are evolving in photovoltaic fault management systems. Until recently, model-based technique, output signal analysis technique, statistically based technique, and machine learning techniques are the four main advanced fault detection methods that researchers have widely studied. This study has identified the limitations and advantages of previous photovoltaic fault detection and monitoring techniques, especially their applicability to all sizes of photovoltaic systems. This study proposes a multi-scale dual-stage photovoltaic fault detection and monitoring technique for better system safety, efficiency, and reliability. Challenges and suggestions for future research directions are also provided in this study. Overall, this study shall provide researchers and policymakers with a valuable reference for developing better fault detection and monitoring techniques for photovoltaic systems.

***Index Terms*—**Advanced fault detection and monitoring techniques, machine learning techniques, model-based technique, output signal analysis technique, photovoltaic, statistically based technique

I. INTRODUCTION

The power generation from the solar photovoltaics (PV) system has experienced substantial increases in the last decade worldwide [1]. However, along with this growth, the associated risks also increased significantly. Common faults in the PV system can be caused by physical or component damage, electrical and mechanical damage, defects in system design, environmental impacts, improper installation, and unscheduled maintenance [2,3]. Therefore, a PV fault monitoring system is crucial in increasing PV systems' reliability, efficiency, and safety.

Conventional protection devices (CPD) such as fuses and circuit breakers are normally used as protection devices because they are easy to handle and inexpensive. The CPD can detect faults and isolate faulty circuits only for large fault currents [4]. Studies have also proven that CPDs are usually hard to detect faults if the voltage difference is minor, such as faults under low irradiance transition conditions that occur during partial shading, degradation, and the transition of night-to-day [5,6]. As a result, undetected faults are hidden, reduce the system's efficiency and reliability, and may cause severe effects such as electric shock or fire hazards [5,7-9]. Hence, advanced fault detection and monitoring technique (FDMT) is needed for maintaining PV system, identifying the root cause of failures, and recommending corrective action for the PV system to operate or function accurately.

Researchers have conducted many studies on advanced FDMT to develop better PV fault detection and monitoring systems, especially to detect potential faults on the DC side, which is the dominant part of fire risk [8]. According to [10], four main FDMTs are commonly used in PV systems: the model-based power loss and I, V curve approaches are the most widely used, followed by machine learning and statistically based techniques (SBTs). The least used is the output signal analysis technique.

The model-based approach compares the expected data obtained from the simulation process with data measured from an experiment or data collected from a real PV system [11,17], while today's most favorable method, the machine learning technique (MLT), which mainly consists

Corresponding author:
Siti Nor Azlina Mohd Ghazali

E-mail:
pes19001@stdmail.ump.edu.my

Received: March 2, 2022

Revised: May 21, 2022

Accepted: July 5, 2022

Publication Date: September 15, 2022

DOI: 10.5152/electrica.2022.22024

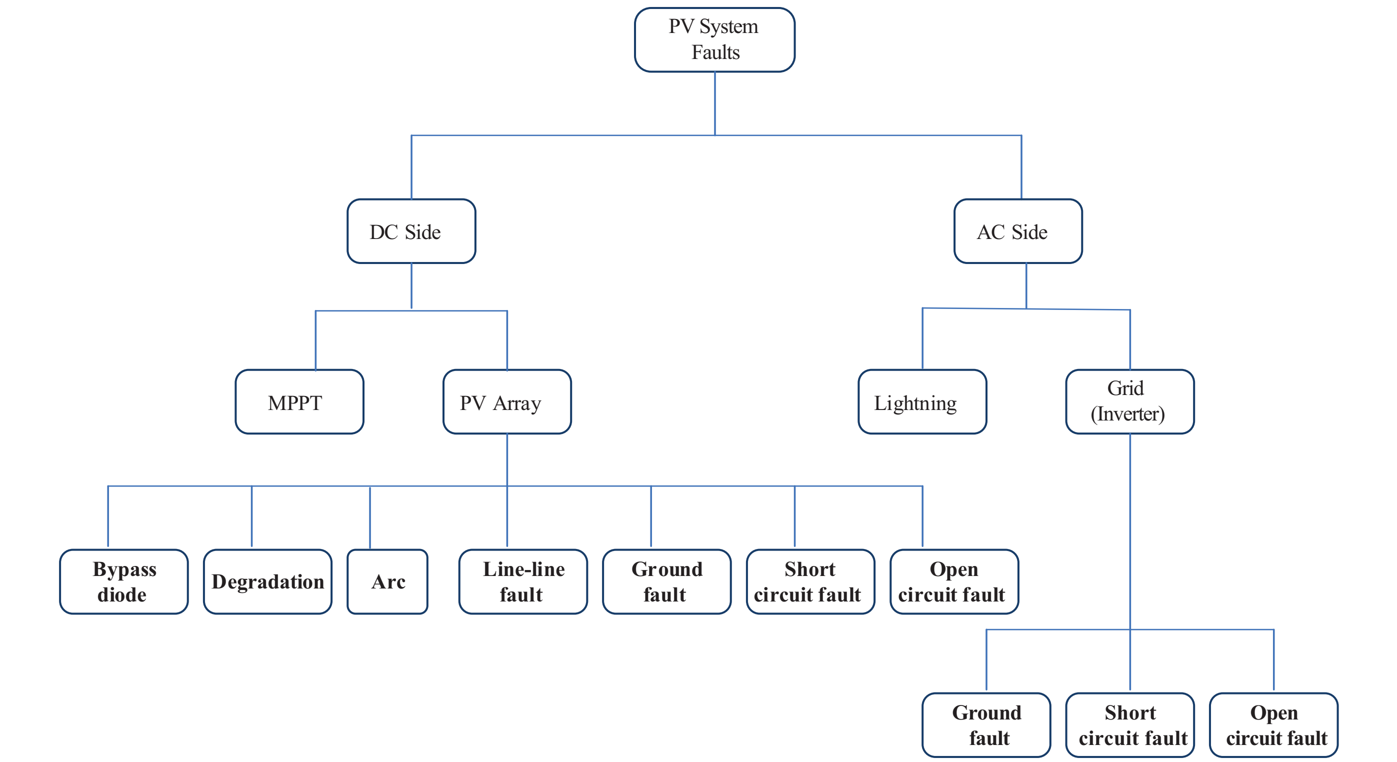

Fig. 1. Common faults on the DC and AC sides of the PV system.

of three main algorithm approaches, supervised learning, semi-supervised learning, and unsupervised learning, exploits artificial intelligence (AI) in completing the task [18-25]. Meanwhile, in the statistical analysis approaches, the most common application is determining the threshold value of each monitored parameter and comparing it with the measured value of the threshold limit (lower

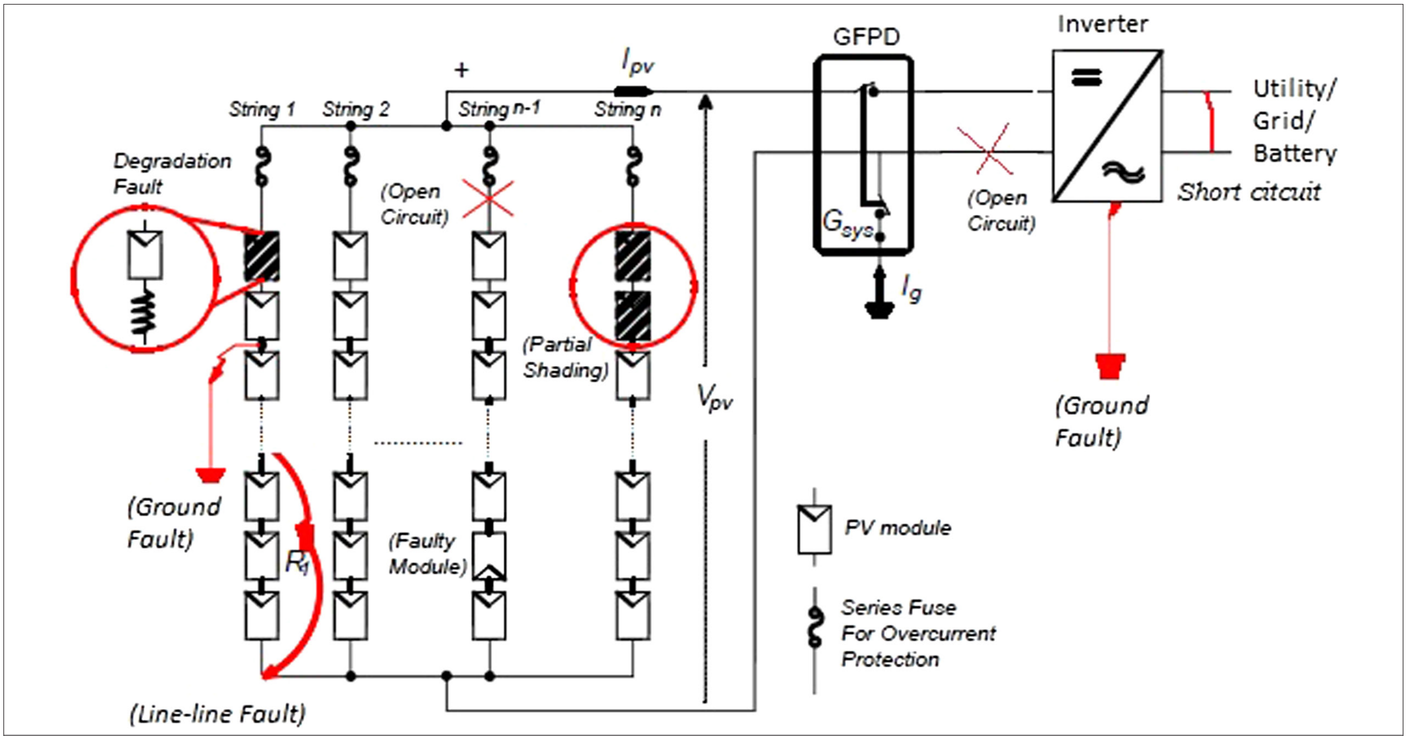

Fig. 2. Brief description of common faults on DC and AC sides of the PV system.

and upper limit) to obtain normal or faulty conditions of the PV system [17,26-30]. Finally, the output signal analysis works on the frequency time-domain examination to detect any abnormalities in the sample while identifying faults in the PV system [31-36].

Many studies have analyzed the effectiveness of these four advanced FDMTs mainly on the accuracy, ability to detect and diagnose faults, the complexity of integration with PV systems, and implementation costs. Nevertheless, this study found that the development of these fault detection techniques still has gaps that can be improved, especially for the fault localization of large-scale PV systems, which is always challenging and time-consuming [37,38], the practicability for PV maintenance work [39], and suitability for all PV array sizes. Hence, an improved version of the fault detection and monitoring approach has been proposed in this study. This study is organized as follows: first, this study describes common faults in PV systems. Then it analyzes the main advanced FDMTs established from previous studies to find their advantages and limitations. Subsequently, this study provides recommendations and discussions to develop an improved PV system fault detection and monitoring model. Finally, this study presents conclusions and discusses the work for the future direction.

II. COMMON FAULTS IN THE PV SYSTEM

Common faults on the DC and AC sides of PV systems can be categorized according to time characteristics as permanent, intermittent, and incipient [37]. Permanent faults remain in place until corrections are made, such as line–line faults, open-circuit faults, ground faults, and arc faults. Intermittent faults refer to temporary effects such as shading, leaves, bird drops, dust, and snow accumulation. On the other hand, incipient fault can occur due to cell degradation, corrosion, and partial damage to the joints. Untreated incipient faults may lead to permanent faults. The common fault classifications for PV systems are illustrated in Fig. 1, with a brief description of those faults shown in Fig. 2 [37,40,41]. Meanwhile, a brief overview of deterministic features of the faults, their potential causes, and effects are depicted in Table I [3,8,37,42].

III. OVERVIEW OF MAIN ADVANCED PV FAULT DETECTION TECHNIQUE

As CPD cannot detect all the defects and failures in PV systems, many advanced FDMTs have been developed for reliable PV system protection methods. This study evaluates model-based power loss and IV curve analysis, MLT, statistically based analysis, and output signal analysis proposed in the previous studies regarding their reliability and feasibility.

A. Model-based Technique

Model-based Technique (MBT) uses the principle of comparing simulation data (expected values) with experimental or measured data (output) from a real system to find the abnormalities. The common procedure of MBT as fault detection and diagnosis technique can be seen in Fig. 3 [11-13,15,16,43]. The basic formula for the MBT involving a threshold that describes the limits of the PV system operational status detection variable is given in (1). In contrast, the MBT involves a threshold limit in (2).

$$\begin{cases} Data_{meas} = xData_{sim} \\ = Status; \, If\ Yes\left(Normal\right) If\ No\left(Faulty\right) \end{cases} \tag{1}$$

TABLE I. THE COMMON FAULTS IN A PV SYSTEM, THEIR POTENTIAL CAUSES, AND EFFECTS

Fault Type	Causes of Fault/Defect	Potential Effect
Short-circuit fault/ line–line fault	Undersize overcurrent protection devices, undersize cable/wire, poor installation practices, or human error	Reduce power output, overheating, wire insulation damage
Open-circuit fault	Disconnected terminal circuit	Reduce power output
Mismatch/partial shading	Environmental effects, uneven radiation received by PV array, inadequate maintenance	Reduce power output
Ground fault	Incorrect grounding system design, poor installation practices, or human error	Tripping/no power output
Arc fault/ hot spot	Poor installation practices or human error	Fire risk
Lightning strike	Natural hazard	Fire risk
Degradation	Weathering and aging effect	Reduce power generation
Reverse bypass diode	Poor installation practices or human error	No power output

PV, photovoltaic.

$$\begin{cases} Data_{\_thmax} \le Data_{sim} \le Data_{\_thmin} \\ = Status; \, If\ Yes\left(normal\right), If\ No\left(Faulty\right) \end{cases} \tag{2}$$

From (1), x is the tolerance factor in avoiding tripping. While in (2), *thmax* and thmin refer to the upper and lower thresholds, respectively.

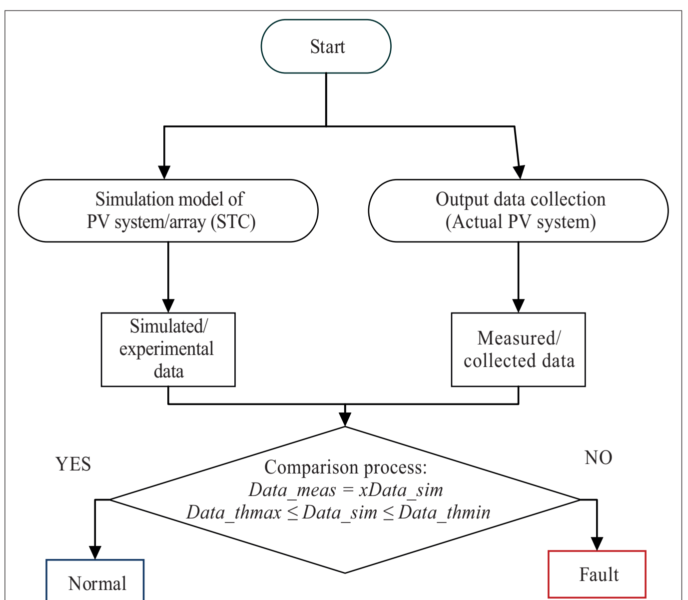

Fig. 3. The common procedure of MBT.

Studies [11,13] compared**,** measured, and simulated output ratios, including I–V curve signature analysis, module degradation factor, and power performance, to develop a real-time FDMT algorithm. This algorithm has successfully detected partial shading, faulty module, degradation, ground fault, line–line fault, and open-circuit fault. Another real-time FDMT algorithm was developed [14]. This FDMT was formulated based on theoretical voltage and power ratio analysis for fault detection and hybrid with fuzzy logic (FL) algorithm to detect partial shading faults on the DC side of the grid-connected PV system (GCPVS).

A study [15] used estimated and measured maximum power point current and maximum power values to detect, diagnose, and locate a line–line fault, ground fault, and partial shading for PV arrays with blocking diodes. On the contrary, a study by [16] proposed a fault detection algorithm using artificial bee colonies (ABC) as an optimizer. This study has successfully detected four faults using estimated and measured values of maximum power point voltage, maximum power point current, and maximum power. Furthermore [17], used Gray wolf optimization (GWO) to detect, diagnose, and locate open-circuit and short-circuit faults at the string level of PV arrays. A detailed summary of some featured studies using the MBT is summarized in Table II.

The table shows that most of the proposed MBTs were usually developed through simulation processes and experimental procedures involving weather stations to obtain actual or measured data. The installation method was easily executed, required medium to high installation costs, and obtained medium to high accuracy. Nevertheless, most were developed to detect specific faults and were not tested for fault locations. In addition, almost all of these studies were not tested or evaluated for large-scale PV arrays.

B. Machine Learning Technique

MLT consists of various methods, principles, and structures. They require input data and parameters from weather and meteorological station. The MLT has three main algorithms: supervised learning, semi-supervised learning, and unsupervised learning, which utilize computer intellect to complete the task [44]. The supervised learning process uses training data that are fully labeled with a known class label and is the most used MLT algorithm, such as an artificial neural network (ANN), which is normally used for classification. The common process of MLT as fault detection and diagnosis technique is depicted in Fig. 4 [18-25].

The study [18] employed the ANN and an optimized Kernel extreme learning machine for fault diagnosis and verified high accuracy.

TABLE II. A DETAILED SUMMARY OF SOME STUDIES USING THE MODEL-BASED TECHNIQUE

Ref	Description of Study (Approaches and Type of PV System)	Software/ Programming	Tested/Evaluated (PV Array Scale)			Validation	Cost of Installation	Accuracy Achieved
			Small	Medium	Big			
[11]	• Fault detection: I-V curve real-time analysis • Stand alone • Fault detected: PS on DC side • Easy integration with PV array/system	LabVIEW	√			Simulation and experiments	Medium	Medium
[13]	• Fault detection: module degradation factor • Stand alone • Fault detected: DA and LLF on the DC side. • Easy integration with PV array/system	-	√	√		Experiments	High	Medium
[14]	• Fault detection: VR and PR • Fault diagnosis: FL • Grid-tied • Fault detected: PS on DC side • Medium complexity of integration with PV array/system	LabVIEW	√	√		Simulation and experiments	Medium to high	High
[15]	• Fault detection and diagnosis: Voc, Isc, Pmpp, and Impp • Stand alone with MPPT • Fault detected: LLF, LGF, and PS on the DC side. • Medium complexity of integration with PV array/system	PSIM	√			Simulation and experiments	High	High
[16]	• Fault detection and diagnosis: Vmpp, Impp, Pmpp, and ABC algorithm • Grid-tied with MPPT • Fault detected: OCF, SCF, PS, and IF on DC side and AC side. • Medium complexity of integration with PV array/system	MATLAB/PSIM	√			Simulation and experiments	Medium to high	High
[17]	• Fault detection: Output current • Fault diagnosis: GWO • Stand alone with MPPT • Fault detected: OCF and SCF on the DC side. • Medium complexity of integration with PV array/system	MATLAB	√			Simulation and experiments	Medium to high	High

VR, voltage ratio; PR, power ratio; FL, fuzzy logic; LLF, line–line fault; LGF, line–ground fault; PS, partial shading; OCF, open-circuit fault; SCF, short-circuit fault; IF, inverter fault; FM, faulty module; GF, ground fault; DA, degradation array; IF, inverter fault; Voc, open circuit voltage; Isc, short circuit current; Pmpp, maximum power point; Impp, maximum power point current; Vmpp, maximum power point voltage, MPPT, maximum power point tracking.

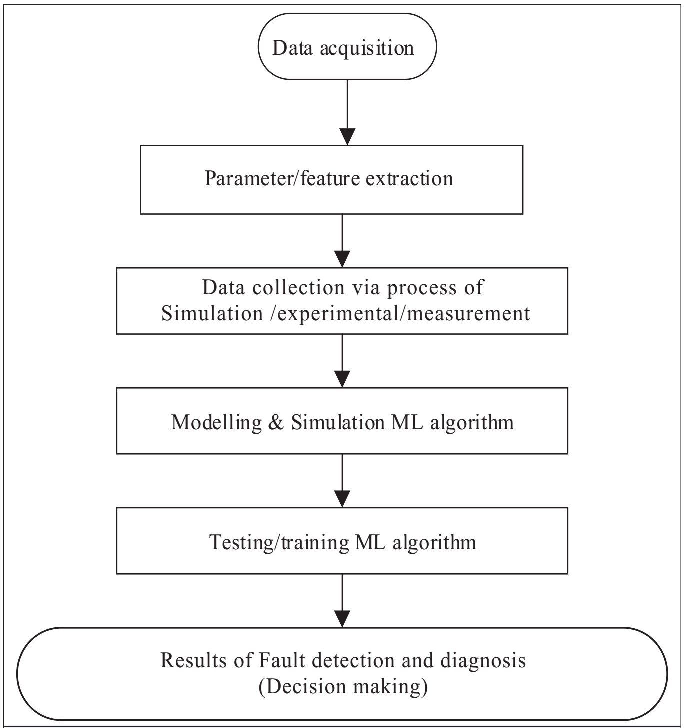

Fig. 4. The common procedure of MLT.

[45] has studied the combination of radial basis function networks (RBFN) and the wavelet transform in PV fault classification. The RBFN is a feed-forward neural network, and a faster learning speed has performed better than a conventional classifier.

Besides that [19] developed a fault detection algorithm for a line–line fault in a PV array based on a support vector machine (SVM) classifier. Meanwhile [20] explored random forest-based (RF) fault diagnosis for PV arrays using array voltage and string current. Lastly [21] has proven that the MLT using K-nearest neighbors has successfully detected, diagnosed, and localized electrical faults and partial shading.

Semi-supervised learning is a combination of supervised and unsupervised learning approaches. It uses both labeled and unlabeled data for training. This technique learns from mistakes to make a decision, such as in [22], which used a graph-based semi-supervised learning procedure. Lastly, unsupervised learning uses only unlabeled training data for processing tasks of clustering or prediction, such as in [23] using FL. It has successfully detected and diagnosed faults on the DC side of GCPVS. A detailed summary of these studies is presented in Table III.

The table shows that most proposed MLT detection methods were developed through simulation and experimental procedures involving weather stations to obtain data. These methods have achieved high accuracy in detecting and diagnosing faults. However, most

TABLE III. A DETAILED SUMMARY OF SOME FEATURED STUDIES USING THE MACHINE LEARNING TECHNIQUES

Ref	Description of Study (Approaches and Type of PV System)	Software/ Programming	Tested/Evaluated (PV Array Scale) Small	Medium	Big	Validation	Cost of Installation	Accuracy Achieved
[18]	• Fault detection: ANN • Fault diagnosis: KELM • Stand alone • Fault detected: DA, SCF, OCF, and PS on the DC side	MATLAB Simulink	√	√		Simulation and experiments	Medium	High
[19]	• Fault detection: SVM • Stand alone • Fault detected: LLF, DA, and PS on the DC side	MATLAB Simulink	√			Simulation and experiments	Medium	High
[20]	• Fault detection and diagnosis: RF • Grid-tied • Fault detected: DA, PS, and LLF on DC side and AC side	MATLAB Simulink	√			Simulation and experiments	High	Medium
[21]	• Fault detection, diagnosis, and location: KNN • Grid-tied • Fault detected: OCF, LLF, and PS on DC side and AC side.	MATLAB/ PSIM	√	√		Simulation and experiments	Medium to high	Medium
[22]	• Fault detection and diagnosis: GBSL • Stand alone • Fault detected: LLF and OCF on the DC side.	MATLAB Simulink	√			Simulation and experiments	High	Medium
[23]	• Fault detection: ANN and FL • Grid-tied • Fault detected: OCF and SCF on the DC side.	MATLAB Simulink	√			Simulation and experiments	Medium	High
[45]	• Fault classification: RBFN combined with a wavelet-based approach for feature extraction • Stand alone • Fault detected: PS, LLF, SCF, OCF, PS, and FM on DC side and AC side	MATLAB Simulink	√			Simulation and experiments	Low	High

LLF, line–, line fault; LGF, line, ground fault; PS, partial shading; OCF, open, circuit fault; SCF, short, circuit fault; IF, inverter fault; FM, faulty module; GF, ground fault; DA, degradation array; IF, inverter fault; ANN, artificial neural network; KELM, Kernel extreme learning machine; SVM, support vector machine; RF, random forest; KNN, K-nearest neighbors; GBSL, graph-based semi-supervised learning; RBFN, radial basis function networks.

studies did not offer the fault location. In addition, they required moderate to high implementation costs due to the need for specialized computer software and specialized skills to execute the procedure. In addition, all of these MLTs were not tested or evaluated on large-scale PV arrays.

C. Statistically based Technique

SBT uses a data-based approach to build a model. Usually, the SBT approach applies a difference in the mean or variance to indicate any errors in the system [46]. Figure 5 describes the common process of SBT in detecting and diagnosing faults in the PV system [17,26-29].

A study by [26] proposed a *t*, test analysis of power and voltage ratio to determine a faulty PV module, faulty string, faulty bypass diode, and faulty MPPT on the DC and AC sides of the GCPVS, while [27] introduced the PV fault detection method with better noise robustness and monitoring quality using the independently generalized likelihood ratio test.

A study by [28] developed a fault monitoring method using an exponentially weighted moving average (EWMA). The technique successfully detected open-circuit fault, short-circuit fault, and partial shading in the PV system. To further explore this technique [29] presented univariate and multivariate exponentially weighted moving average charts to monitor the performance of the PV system. The

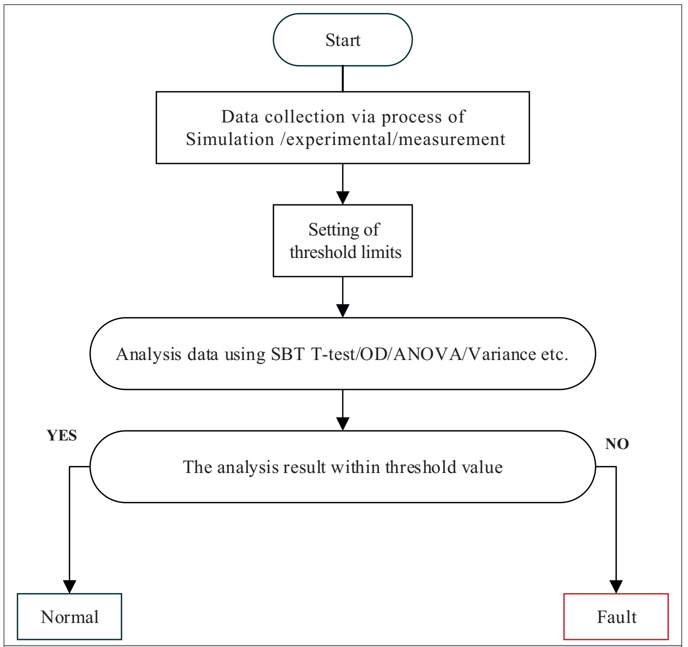

Fig. 5. The common procedure of SBT.

TABLE IV. A DETAILED SUMMARY OF SOME FEATURED STUDIES USING THE STATISTICALLY, BASED TECHNIQUES

Ref	Description of Study (Approaches and Type of PV System)	Software/ Programming	Tested/Evaluated (PV Array Scale) Small	Medium	Big	Validation	Cost of Installation	Accuracy Achieved
[26]	• Fault detection and diagnosis: *t*-test statistical analysis with threshold limits on PR and VR • Grid-tied • Fault detected: FM and PS on DC side and AC side.	LabVIEW	√	√		Simulation and experiments	Medium to high	High
[27]	• Fault detection and diagnosis: GLRT with threshold limits • Grid-tied • Fault detected: PS and FM on the DC side	MATLAB Simulink	√			Simulation and experiments	Medium	Medium
[28]	• Fault detection and diagnosis: EWMA with threshold limits • Grid-tied • Fault detected: temporary PS, OCF, and SCF on the DC side	MATLAB/ PSIM	√	√		Simulation and experiments	Medium to high	Medium
[29]	• Fault detection and diagnosis: MEWMA with threshold limits • Grid-tied • Fault detected: Temporary PS, OCF, SCF, and DA on the DC side	MATLAB/ PSIM	√			Simulation and experiments	High	High
[30]	• Fault detection and diagnosis using VPCA • Grid-tied • Fault detected: PS on DC side	VPCA computational model	√	√		Simulation and experiments	High	High
[47]	• Fault detection, diagnosis, and location: OD and regression coefficient extraction • Grid-tied • Fault detected: LLF on DC side	MATLAB Simulink	√	√		Simulation and experiments	Medium to high	High

VR, voltage ratio; PR, power ratio; LLF, line–line fault; PS, partial shading; OCF, open-circuit fault; SCF, short-circuit fault; IF, inverter fault; FM, faulty module; GF, ground fault; DA, degradation array; IF, inverter fault; GLRT, generalized likelihood ratio test; EWMA, exponentially weighted moving average; MEWMA, univariate and multivariate exponentially weighted moving average; VPCA, vertical principal component analysis; OD, outlier detection.

study showed that the proposed method could detect and distinguish between permanent faults (short-circuit fault, open-circuit fault, and degradation) and temporary faults (a small portion of partial shading). Another interesting study by [47] developed fault detection and location using residual-based outlier detection and regression coefficient extraction governed by the PV system's predefined threshold, producing better performance. On the other hand [30], presented the partial shading fault detection method using vertical principal component analysis, which has proven to obtain better results than the standard PCA method. A detailed review of these techniques is presented in Table IV.

The table shows that SBT detection methods were developed through simulation processes and experimental procedures involving weather stations to obtain data. This technique used a threshold limit for determining the normal or faulty condition of the PV system and has achieved high accuracy. They required medium to high installation costs. However, most of these studies did not offer fault locations and were not tested or evaluated on large-scale PV arrays.

D. Output Signal Analysis Technique

Generally, the output signal analysis technique (OST) works on the frequency/time domain analysis to detect any changes in the sample during the fault detection process. A time-domain signal is measured by varying the amplitude of the signal layers over time. While a frequency-domain signal is converted into a time-domain signal to noise, indicating fault detection and diagnosis errors. The frequency-domain signal can emit the sound found in the time domain signal. Lastly, the time-frequency signal works in both the time and frequency simultaneously and is applied when the signals in short-term vary significantly during the process. The common procedure of OST for PV fault detection can be seen in Fig. 6 [31-36].

In [31], fast fourier transform was used for arc fault detection on the DC and AC sides. [32] developed a DC arc-fault detection method using discrete wavelet transform (DWT) to analyze the spectral characteristics of electromagnetic radiation signals corresponding to abrupt changes in arcing current. Temporarily [33] applied wavelet packets transform (WPT) to develop an online method to identify PV array faults under low irradiation. Their study has proven that the performance of the WPT method is better than DWT in detecting failures.

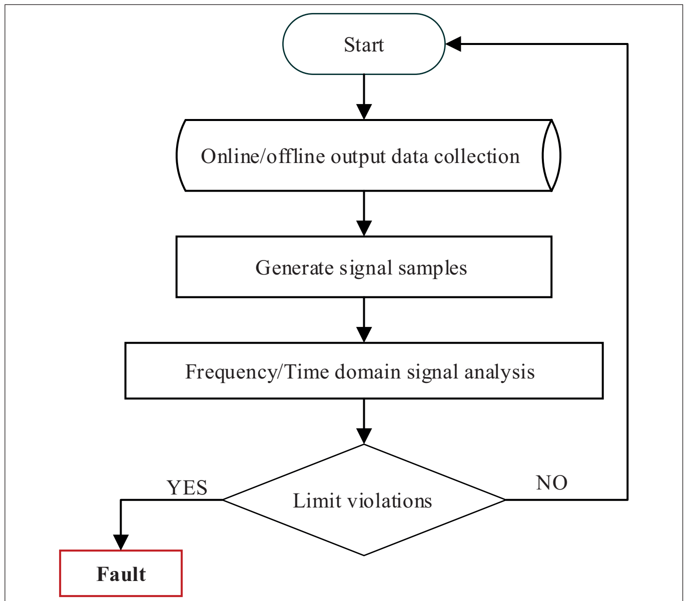

Fig. 6. The common procedure of OST.

OST studies based on spread-spectrum time-domain reflectometry have attracted many researchers. For example, [34] developed this technique for detecting and locating disconnection faults in PV

TABLE V. A DETAILED SUMMARY OF SOME FEATURED STUDIES USING THE OUTPUT SIGNAL ANALYSIS TECHNIQUES

Ref	Description of Study (Approaches and Type of PV System)	Software/ Programming	Tested/Evaluated (PV Array Scale) Small	Medium	Big	Validation	Cost of Installation	Accuracy Achieved
[31]	• Fault detection: FFT • Stand alone • Fault detected: Arc fault on DC side and AC side.	-	√			Experiments	High	High
[32]	• Fault detection: DWT • Stand alone • Fault detected: Arc fault on DC side	-	√			Experiments	High	High
[33]	• Fault detection: WPT • Grid-tied • Fault detected: LLF, GF, and PS on the DC side	MATLAB Simulink	√			Simulation and experiments	High	High
[34]	• Fault detection: SSTDR • Stand alone • Fault detected: OCF and SCF in PV string	-	√			Experiments	Medium	Medium
[35]	• Fault detection: SSTDR • Stand alone • Fault detected: analysis of faults of complex loads.	MATLAB Simulink	√			Simulation and experiments	High	High

LLF, line–line fault; LGF, line–ground fault; PS, partial shading; OCF, open–circuit fault; SCF, short–circuit fault; IF, inverter fault; FM, faulty module; GF, ground fault; DA, degradation array; IF, inverter fault; FFT, fast fourier transform; DWT, discrete wavelet transform, WPT, wavelet packets transform; SSTDR, spread-spectrum time-domain reflectometry.

power plants, whereas [35] used this technique for investigating faults in PV systems due to complex loads. A detailed review of these approaches is summarized in Table V.

The table shows that the OST detection method accurately detects PV system/array failures. However, it involved high costs because it required sophisticated tools to produce high-speed sampling. Also, none of these methods offered a fault location. Finally, almost all studies were tested or evaluated only on small-scale PV arrays/systems.

IV. DISCUSSION AND RECOMMENDATION

Table VI presents an analytical evaluation showing that the methods/techniques developed have achieved good accuracy and contributed to new knowledge. However, most were only developed for the specific fault(s) and did not provide fault locations. In addition, the requirements of data loggers, specialized computer software, and specialized skills to operate the software in collecting data needed substantial costs. Furthermore, no assessment of the suitability of PV system maintenance was performed. Finally, almost all those studies were only tested and evaluated on small-scale PV systems. Consequently, their practicality for medium and large-scale PV systems cannot be determined.

Hence, this study proposes a multi-scale dual-stage (MsDs) PV FDMT and the main contributions as the following:

a. MsDs is a low-cost and easy execution model to detect, diagnose, and locate a variety of common faults on PV arrays without the interruption of system operation.
b. MsDs is an AI-based FDMT approach that can benefit PV system maintenance work. MsDs can distinguish different incidents or failures with different characteristics that require

TABLE VI. ANALYTICAL EVALUATION OF MAIN ADVANCED FDMT

Ref	Description of study	Summary results/limitations
[15]	The circuit-based analysis for detecting and diagnosing faults compared Voc, Isc, Pmpp, and Impp to detect LLF, LGF, and PS on the DC side of the PV array for a stand-alone PV system with MPPT. The proposed method is viable for the scalable PV array size.	The algorithm developed through simulation and experiment has obtained good results and involved moderate costs. The development cost of this algorithm will increase as the scale size of the PV array experiment setup increases. The procedure was quite complicated during data acquisition involving various equipment for circuit analysis. Furthermore, this study was not tested/evaluated on medium and large-scale PV systems.
[16]	The fault detection and diagnosis are based on a model-based technique using Vmpp, Impp, Pmpp, and ABC algorithms to detect OCF, SCF, PS, and IF on the DC and AC sides GCPVS with MPPT.	This study has proven achieved good accuracy and detected multiple faults. The proposed technique using upper and lower thresholds for each indicator of good or faulty PV system operation may affect the study's accuracy as it is difficult to obtain an accurate and appropriate threshold. In addition, this study was not tested/evaluated on medium and large-scale PV systems.
[18]	The fault diagnosis modeling using KELM combined with Nelder–Mead Simplex to optimize the parameter to detect DA, SCF, OCF, and PS on the stand-alone PV system's DC side.	For the labeled fault data sample, KELM, which has a very fast learning speed and good generalization performance, has produced high accuracy results in this study. However, in reality, labeled data are not always available, and preparing/generating labeled data samples through fault simulations and field experiments will cost high.
[30]	The procedure for detecting PS on the DC side of the GCPVS system using VPCA. The study also extended the contribution of a plot method for diagnosing the faults to the interval case.	This proposed VPCA procedure, which considered the uncertainties of the PV system and used Q-statistics and T2-statistics, has been shown to obtain better accuracy than conventional PCA. Nevertheless, this proposed method has not tested its feasibility on medium and large-scale PV systems.
[33]	The online fault detection method used WPT to distinguish PS from the faulty conditions on the DC part of CGPVS with MPPT. The array voltage and current were selected as inputs, and threshold values were generated from the MATLAB/Simulink simulations.	The proposed method succeeded in distinguishing PS from other faults because WPT can detect abnormalities in the signal (data) that undergoes abrupt changes. This method is useful for avoiding false alarm tripping for low mismatch fault percentages and low irradiation conditions. However, wrongly determining the pre-set threshold can affect the accuracy, and this study only tested on small-scale PV arrays.
[34]	The study was to find and locate OCF and SCF in PV strings using STTDR for the stand-alone PV systems with complex (capacitive) loads. This technique was developed by the experimental setup of uneven PV panel configuration.	OCF and SCF were successfully detected when changes in impedance due to PV module discontinuities were analyzed using the SSTDR technique. However, when more panels were connected, which caused an increase in impedance (capacitance), finding the location of the faults became more difficult due to the more peaks, and difficult to analyze. Moreover, this study was only tested on small-scale PV arrays.
[45]	The fault classification model using RBFN combined with a wavelet-based approach for feature extraction to detect PS, LLF, SCF, OCF, PS, and FM on the DC side and the AC side for stand-alone PV systems.	The performance of the developed technique was proven better than conventional methods with a very short detection time. The technique was tested on a 1 kW single-phase stand-alone PV system, which described 100% training efficiency under 13 seconds and 97% test efficiency under 0.2 seconds. However, the performance of this proposed approach has not been tested/evaluated on medium and large-scale PV systems, which are expected to require more simulation time.
[47]	The approach of the outlier rule and comparison of a pre-set threshold for current string measurements to detect and identify the location of wiring faults either on the intra-string or cross-string for the DC side of the GCPVS.	A combination of analytical and regression expressions that depend on the type of fault, radiation level, and string current measurements have successfully detected and located LLFs on intra -string/cross-string. Obtaining labeled data to validate the proposed technique involved a lot of equipment and high costs. Furthermore, wrongly determining the pre-set threshold can affect the accuracy, and this study was not tested/evaluated on medium and large-scale PV systems.

specialize competent people and different tools and techniques to address and implement corrections [48]. Thus, it can reduce response time, reduce the cost of corrective work, and optimize the system operation.

c. MsDs is recommended to be tested and evaluated on small, medium, and large PV array models to ensure its feasibility on all scales of PV arrays.

A. Multi-scale Dual-stage Photovoltaic Fault Detection and Monitoring Technique

Figure 7 depicts the proposed MsDs flow chart, consisting of a hybrid algorithm, where stage-1 applies a comparison-based algorithm for fault detection procedures and stage-2 employs MLTs for fault diagnosis and location procedures. The proposed fault detection algorithm is developed to detect six common faults on the DC side of the PV arrays.

For the data acquisition, the study proposed to develop a PV array fault-free model (*PV_nofault*) and a PV array fault model (*PV_fault*^{n}) for partial shading, ground fault, line–line fault, open-circuit fault, degradation array, and faulty module. Where $n = 1–6$ represents these six faults. The PV array modeling is developed using MATLAB Simulink, a modified version [18]. The values for the input parameters, Isc, Voc, Iph, N, Rs, and Rsh, are obtained from the manufacturer's PV module specification data. Fault detection at stage 1 is based on a comparison between the PV yield by the model *PV_nofault* and *PV_fault*^{n} in detecting abnormalities. The yield of solar panel, *r* (kWp), is formulated as in (3). It is expected that *r* from the *PV_nofault* to be higher than *r* from the *PV_fault*^{n} [49]. The fault detection condition is given in (4)

$$r = \frac{Electric\ power\ (kWp)}{area\ of\ one\ panel} \tag{3}$$

$$r_{PV_{nofault}} \geq r_{PV_{fault}} = \text{If YES, fault detected; f NO, normal condition} \tag{4}$$

Stage 2 is fault diagnosis and location procedures to identify whether the fault is located at a module, string, or array. The training at stage 2 is suggested to be performed with cross-validation together

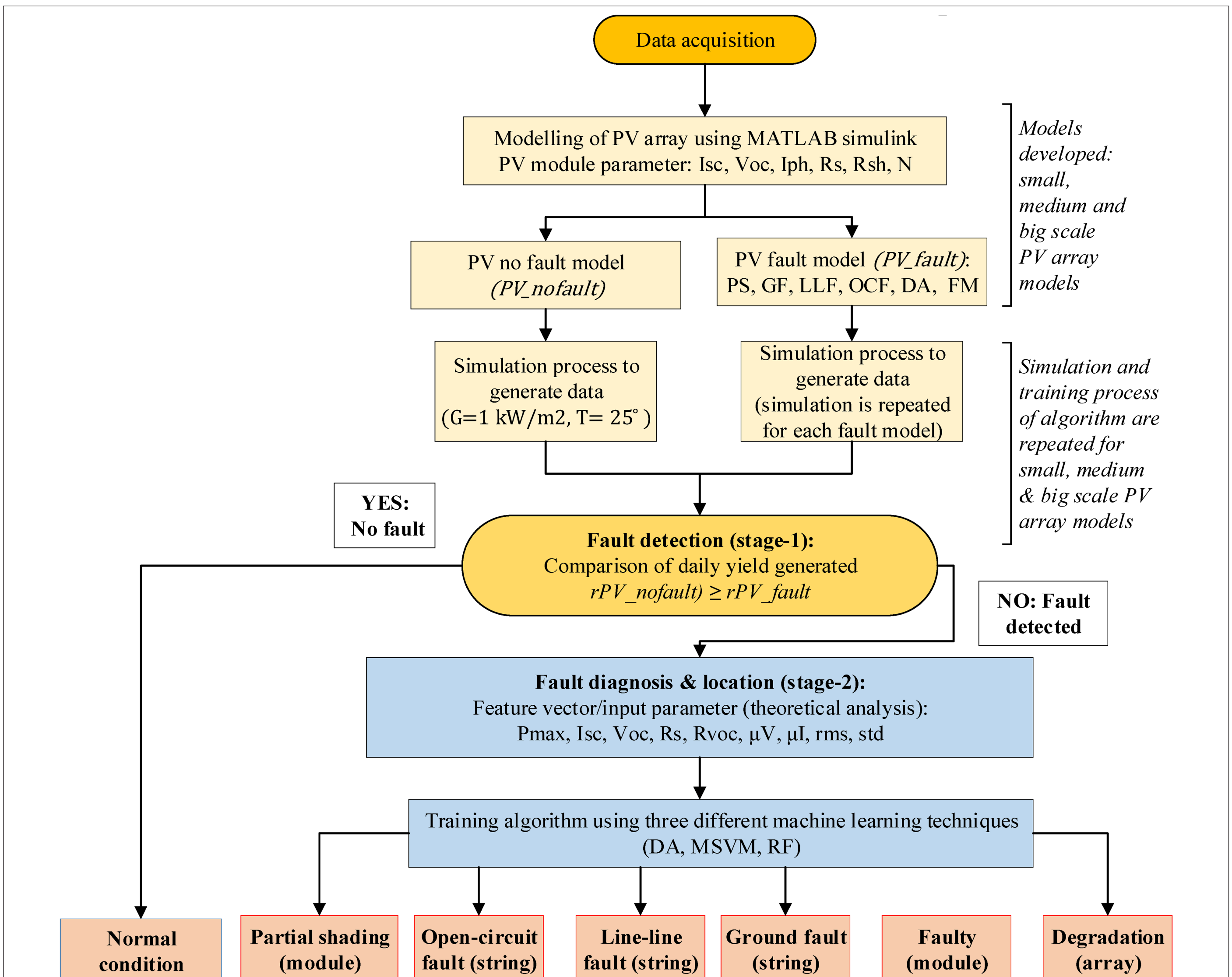

Fig. 7. Flowchart of the proposed multi-scale dual-stage PV fault detection and monitoring technique.

with three different types of MLT algorithms: discrimination analysis (DA), RF, and multi-class support vector machine, to obtain the best classification accuracy. The input parameters are the feature vectors of maximum power, short-circuit current, open-circuit voltage, mean current and voltage, and the standard deviation. These input parameters are selected based on the theoretical analysis of I–V curves generated from the PV array modeling simulation process.

B. Performance Evaluation of the Proposed MsDs Technique

To validate the performance of MsDs as feasible to all PV array scales and practical for PV system maintenance, this study recommends:

1. to develop and simulate small, medium, and large-scale PV array fault-free and PV array fault models.
2. to train/test the proposed algorithm using several MLTs to acquire the best accuracy results.
3. to train/test the proposed algorithm using labeled and unlabeled data collected from actual operating PV systems to evaluate its practicality for PV maintenance.

C. Challenges and Improvements for Future

This study captures some challenging issues; some modifications are recommended to be implemented for future work are proposed as follows:

1. It is necessary to demonstrate the performance of the proposed method/technique by training/testing using the actual data collected, as the accuracy of the results will usually be reduced compared to the simulated data [50]. However, controlling external factors during the experimental setup is challenging.
2. The fault detection procedure for PV maintenance work uses unlabeled data from the real operating PV system to find hidden PV fault(s). However, it is challenging to collect fault sample data and control external factors of operating conditions and the external temperature of PV arrays for evaluating the proposed algorithms for accurate and effective PV system FDMT.

V. CONCLUSION

Studies on developing advanced fault detection methods/techniques for PV arrays/systems have been widely conducted and proven that it can detect and diagnose faults well. Developing a good fault detection technique is important and indispensable since conventional protection devices such as circuit breakers and fuses cannot clean up all PV system failures due to PV's non-linear output characteristics. Untreated faults will be hidden in the system and can reduce the system's productivity, further may lead to the risk of fire, which can cause loss of life and destruction to the system and buildings.

Previously, many reviews have been performed on PV system advanced fault detection methods/techniques. The purpose is to find the advantages and limitations and seek recommendations to improve the efficiency, reliability, and safety of existing methods/techniques. Nevertheless, this study found gaps that could be improved, especially in terms of suitability in all PV system sizes, including detecting various faults without interruption of system operation and feasibility for PV system maintenance.

This study proposes a MsDs technique for improved fault detection and monitoring of PV systems. This study proposes data acquisition generated from MATLAB Simulink modeling, which requires low cost. The procedure consisted of stage 1 using a comparison-based approach for detecting multiple faults and stage 2 using several machine learning algorithms combined with cross-validation to diagnose and locate detected faults with the best accuracy. Separating the fault detection procedure at stage-1 to check whether fault(s) are hidden before the fault classification procedure at stage 2 is continued is beneficial for PV system maintenance work. In addition, labeled and unlabeled data from the actual operating PV system are recommended to be used in the validation of the proposed algorithm. Subsequently, the developed algorithm is recommended to be tested and evaluated on small, medium, and large-scale PV array modeling to validate its feasibility for multi-scale PV systems.

Overall, the advanced PV fault detection, diagnosis, and location procedures, including recommendations for future work proposed in this study, can be further explored to produce better models and algorithms. Finally, this study shall help PV system researchers and policymakers develop FDMT for better efficiency, reliability, and safety.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – S. N. A. M. G., M. Z. S.; Design – S. N. A. M. G.; Supervision – M. Z. S.; Funding – M. Z. S.; Materials – S. N. A. M. G.; Data Collection and/or Processing – S. N. A. M. G.; Analysis and/or Interpretation – S. N. A. M. G.; Literature Review – S. N. A. M. G.; Writing – S. N. A. M. G.; Critical Review – S. N. A. M. G., M. Z. S.

Declaration of Interests: The authors declare that they have no competing financial interests or personal relationships that could affect the work reported in this article.

Funding: This study received the financial support from Universiti Malaysia Pahang (UMP) under project number PGRS2003189.

REFERENCES

1. BP, "Statistical Review of World Energy, 2020 | 69th Edition," *Bp*, p. 66, 2020.
2. S. Lu, B. T. Phung, and D. Zhang, "A comprehensive review on DC arc faults and their diagnosis methods in photovoltaic systems," *Renew. Sustain. Energy Rev.*, vol. 89, pp. 88–98, 2018. [CrossRef]
3. D. S. Pillai, and N. Rajasekar, "A comprehensive review on protection challenges and fault diagnosis in PV systems," *Renew. Sustain. Energy Rev.*, vol. 91, pp. 18–40, 2018. [CrossRef]
4. S. N. Kamenopoulos, and T. Tsoutsos, "Assessment of the safe operation and maintenance of photovoltaic systems," *Energy*, vol. 93, pp. 1633–1638, 2015. [CrossRef]
5. Y. Zhao, T. Li, X. Zhang, and C. Zhang, "Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future," *Renew. Sustain. Energy Rev.*, vol. 109, pp. 85–101, 2019. [CrossRef]
6. R. Hariharan, M. Chakkarapani, G. Saravana Ilango, and C. Nagamani, "A method to detect photovoltaic array faults and partial shading in PV systems," *IEEE J. Photovolt.*, vol. 6, no. 5, pp. 1278–1285, 2016. [CrossRef]
7. P. Dunbabin, "Fire and Solar PV Systems – Investigations and Evidence", BRE National Solar Centre, UK, Rep. no. P100874-1004, 17th July 2017 [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/630639/fire-solar-pv-systems-investigations-evidence.pdf
8. H. Laukamp, J. Tschupp, F. Ise, and B. Maiworm, "PV Systems and Fire Hazard", Fraunhofer ISE, Jerman, 2018 [Online]. Available: https://www.ise.fraunhofer.de/en/publications/studies/recent-facts-about-pv-in-germany.htm
9. M. D. Vaverkov, J. Winkler, D. Uldrijan, J. Aleksiejuk-gawron, and D. Adamcov, "Fire hazard associated with different types of photovoltaic power plants: Effect of vegetation management," vol. 162, 2021. [CrossRef]

10. A. Livera, M. Theristis, G. Makrides, and G. E. Georghiou, "Recent advances in failure diagnosis techniques based on performance data analysis for grid-connected photovoltaic systems," *Renew. Energy*, vol. 133, pp. 126–143, 2019. [CrossRef]
11. M. H. Ali, A. Rabhi, A. E. El Hajjaji, and G. M. Tina, "Real time fault detection in photovoltaic systems," *Energy Procedia*, vol. 111, pp. 914–923, 2017. [CrossRef]
12. S. Fadhel *et al.*, "PV shading fault detection and classification based on I-V curve using principal component analysis: Application to isolated PV system," *Sol. Energy*, vol. 179, pp. 1–10, 2019. [CrossRef]
13. A. Hazra, S. Das, and M. Basu, "An efficient fault diagnosis method for PV systems following string current," *J. Clean. Prod.*, vol. 154, pp. 220–232, 2017. [CrossRef]
14. M. Dhimish, V. Holmes, B. Mehrdadi, and M. Dales, "Diagnostic method for photovoltaic systems based on six layer detection algorithm," *Electr. Power Syst. Res.*, vol. 151, 26–39, 2017. [CrossRef]
15. A. F. Murtaza, M. Bilal, R. Ahmad, and H. A. Sher, "A circuit analysis-based fault finding algorithm for photovoltaic array under L–L/L-G faults," *IEEE J. Emerg. Sel. Top. Power Electron.*, vol. 8, no. 3, pp. 3067–3076, 2020. [CrossRef]
16. Y. Chaibi, M. Malvoni, A. Chouder, M. Boussetta, and M. Salhi, "Simple and efficient approach to detect and diagnose electrical faults and partial shading in photovoltaic systems," *Energy Convers. Manag.*, vol. 196, 330–343, 2019. [CrossRef]
17. A. Dhoke, R. Sharma, and T. K. Saha, "PV module degradation analysis and impact on settings of overcurrent protection devices," *Sol. Energy*, vol. 160, 360–367, 2018. [CrossRef]
18. Z. Chen, L. Wu, S. Cheng, P. Lin, Y. Wu, and W. Lin, "Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I-V characteristics," *Appl. Energy*, vol. 204, pp. 912–931, 2017. [CrossRef]
19. Z. Yi, and A. H. Etemadi, "Line-to-line fault detection for photovoltaic arrays based on multi-resolution signal decomposition and two-stage support vector machine," *IEEE Trans. Ind. Electron.*, vol. 64, no. 11, 8546–8556, 2017. [CrossRef]
20. Z. Chen *et al.*, "Random forest based intelligent fault diagnosis for PV arrays using array voltage and string currents," *Energy Convers. Manag.*, vol. 178, pp. 250–264, 2018. [CrossRef]
21. S. R. Madeti, and S. N. Singh, "Modeling of PV system based on experimental data for fault detection using kNN method," *Sol. Energy*, vol. 173, 139–151, 2018. [CrossRef]
22. Y. Zhao, R. Ball, J. Mosesian, J. F. De Palma, and B. Lehman, "Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays," *IEEE Trans. Power Electron.*, vol. 30, no. 5, pp. 2848–2858, 2015. [CrossRef]
23. M. Dhimish, V. Holmes, B. Mehrdadi, and M. Dales, "Comparing Mamdani Sugeno fuzzy logic and RBF ANN network for PV fault detection," *Renew. Energy*, vol. 117, pp. 257–274, 2018. [CrossRef]
24. X. Lu *et al.*, "Fault diagnosis for photovoltaic array based on convolutional neural network and electrical time series graph," *Energy Convers. Manag.*, vol. 196, pp. 950–965, 2019. [CrossRef]
25. A. Haque, K. V. S. Bharath, M. A. Khan, I. Khan, and Z. A. Jaffery, "Fault diagnosis of photovoltaic modules," *Energy Sci. Eng.*, vol. 7, no. 3, pp. 622–644, 2019. [CrossRef]
26. M. Dhimish, V. Holmes, and M. Dales, "Parallel fault detection algorithm for grid-connected photovoltaic plants," *Renew. Energy*, vol. 113, pp. 94–111, 2017. [CrossRef]
27. M. Mansouri *et al.*, "An effective statistical fault detection technique for grid connected photovoltaic systems based on an improved generalized likelihood ratio test," *Energy*, vol. 159, pp. 842–856, 2018. [CrossRef]
28. E. Garoudja, F. Harrou, Y. Sun, K. Kara, A. Chouder, and S. Silvestre, "Statistical fault detection in photovoltaic systems," *Sol. Energy*, vol. 150, pp. 485–499, 2017. [CrossRef]
29. F. Harrou, Y. Sun, B. Taghezouit, A. Saidi, and M. E. Hamlati, "Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches," *Renew. Energy*, vol. 116, pp. 22–37, 2018. [CrossRef]
30. L. Rouani, M. F. Harkat, A. Kouadri, and S. Mekhilef, "Shading fault detection in a grid-connected PV system using vertices principal component analysis," *Renew. Energy*, vol. 164, pp. 1527–1539, 2021. [CrossRef]
31. Q. Xiong, R. Xiao, S. Ji, L. Zhu, Y. Liu, and L. Zhong, "Detection method for DC arc based on electromagnetic radiation characteristics," *Gaodianya Jishu/High Volt. Eng.*, vol. 43, no. 9, pp. 2967–2975, 2017. [CrossRef]
32. C. He, L. Mu, and Y. Wang, "The detection of parallel arc fault in photovoltaic systems based on a mixed criterion," *IEEE J. Photovolt.*, vol. 7, no. 6, pp. 1717–1724, 2017. [CrossRef]
33. B. P. Kumar, G. S. Ilango, M. J. B. Reddy, and N. Chilakapati, "Online fault detection and diagnosis in photovoltaic systems using wavelet packets," *IEEE J. Photovolt.*, vol. 8, no. 1, 257–265, 2018. [CrossRef]
34. N. K. T. Jayakumar, M. U. Saleh, E. Benoit, J. Lacombe, M. Scarpulla, and C. Furse, "Fault detection in PV strings using SSTDR," *Usn Radio Sci. Meet. (Joint with AP, S Symp. Usn*, vol. 2018, pp. 9–10, 2018. [CrossRef]
35. E. Benoit *et al.*, "Applicability of SSTDR analysis of complex loads," in Proceedings IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Sci. Meeting, APSURSI 2019, 2019. [CrossRef]
36. S. Roy, M. K. Alam, F. Khan, J. Johnson, and J. Flicker, "An irradiance-independent, robust ground-fault detection scheme for PV arrays based on spread spectrum time-domain reflectometry (SSTDR)," *IEEE Trans. Power Electron.*, vol. 33, no. 8, pp. 7046–7057, 2018. [CrossRef]
37. S. R. Madeti, and S. N. Singh, "A comprehensive study on different types of faults and detection techniques for solar photovoltaic system," *Sol. Energy*, vol. 158, pp. 161–185, 2017. [CrossRef]
38. V. Carletti, A. Greco, A. Saggese, and M. Vento, "An intelligent flying system for automatic detection of faults in photovoltaic plants," *J. Ambient Intell. Human. Comput.*, vol. 11, no. 5, pp. 2027–2040, 2020. [CrossRef]
39. K. Osmani, A. Haddad, T. Lemenand, B. Castanier, and M. Ramadan, "A review on maintenance strategies for PV systems," *Sci. Total Environ.*, vol. 746, p. 141753, 2020. [CrossRef]
40. M. Bouzidi, A. Harrouz, T. Mohammed, and S. Mansouri, "Short and open circuit faults study in the PV system inverter," *IJPEDS*, vol. 12, no. 3, 2021. [CrossRef]
41. R. F. Buzo, H. M. Barradas, and F. B. Leão, "Fault current of PV inverters under grid-connected operation: A review," *J. Control Autom. Electr. Syst.*, vol. 32, no. 4, pp. 1053–1062, 2021. [CrossRef]
42. M. C. Falvo, and S. Capparella, "Safety issues in PV systems: Design choices for a secure fault detection and for preventing fire risk," *Case Stud. Fire Saf.*, vol. 3, pp. 1–16, 2015. [CrossRef]
43. S. Silvestre, *Strategies for Fault Detection and Diagnosis of PV Systems*, Vol. 1. Amsterdam: Elsevier Inc., 2018.
44. A. Mellit, and S. A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," *Prog. Energy Combust. Sci.*, vol. 34, no. 5, pp. 574–632, 2008. [CrossRef]
45. V. S. B. Bharath Kurukuru, F. Blaabjerg, M. A. Khan, and A. Haque, "A novel fault classification approach for photovoltaic systems," *Energies*, vol. 13, no. 2, 2020. [CrossRef]
46. S. Zare, "Fault Detection and Diagnosis of Electric Drives Using Intelligent Machine Learning Approaches", Electronic Theses and Dissertations. University of Windsor, Canada, 2018.
47. A. Dhoke, R. Sharma, and T. K. Saha, "A technique for fault detection, identification and location in solar photovoltaic systems," *Sol. Energy*, vol. 206, pp. 864–874, 2020. [CrossRef]
48. S. Nor, A. Mohd., M. Z. Sujod, and M. S. Jadin, "Forensic of solar PV: A review of potential faults and maintenance strategies," 7th International Conference on Engineering and Emerging Technologies, 27–28 October 2021, Istanbul, Turkey. IEEE, 2021.
49. B. Goss, I. R. Cole, E. Koubli, D. Palmer, T. R. Betts, and R. Gottschalg, *Modelling and Prediction of PV Module Energy Yield*. Amsterdam: Elsevier Ltd., 2017.
50. Y. Emre, K. Vardar, and M. Ali, "Modeling and simulation of PV systems," *IOSR J.E.E.E*, vol. 13, no. 2, pp. 1–11, 2018. [CrossRef]

Siti Nor Azlina Mohd Ghazali was born in Kelantan, Malaysia, in 1978. She received her B.Eng. degree in electrical engineering from the University of Mara Technology, Malaysia, in 2002. Then she received her MSc in Energy Studies from the University of Otago, New Zealand, in 2013. She is a member of the Board of Engineer Malaysia since 2004 and has been appointed as Professional Engineer in April 2018. Then, she has been appointed as ASEAN Charted Professional Engineer in March 2019. She is currently working toward her Ph.D. at the College of Engineering, Department of Electrical and Engineering, Universiti Malaysia Pahang, Malaysia. Her current research interests include PV forensic electrical, PV smart maintenance strategies, and PV smart fault monitoring system.

Muhamad Zahim Bin Sujod was born in Selangor, Malaysia, in 1976. He received the B.Eng. degree and M.Eng. degree in Electrical and Electronics Engineering from the University of Ehime, Ehime, Japan, in 2000 and 2002, respectively, and the Ph.D. degree from the University Duisburg-Essen, Germany, in Power System Engineering, in 2014. He is a member of the Board of Engineer Malaysia (BEM) since January 2004 and has been appointed as Professional Engineer in January 2009. Currently, he is an Associate Professor at the College of Engineering, Universiti Malaysia Pahang, Malaysia. His primary research activities involve renewable energy systems (wind turbine and photovoltaic), energy conversion, energy management, and electrical machines.