

Review

Review of Artificial Intelligence-Based Failure Detection and Diagnosis Methods for Solar Photovoltaic Systems

Ahmad Abubakar * , Carlos Frederico Meschini Almeida and Matheus Gemignani

Department of Electrical Engineering and Automation, Escola Politecnica da Universidade de São Paulo, São Paulo 05508-010, Brazil; cfmalmeida@usp.br (C.F.M.A.); matheusg@alumni.usp.br (M.G.)

* Correspondence: namatoyaa@usp.br; Tel.: +55-11997680626

Abstract: In recent years, the overwhelming growth of solar photovoltaics (PV) energy generation as an alternative to conventional fossil fuel generation has encouraged the search for efficient and more reliable operation and maintenance practices, since PV systems require constant maintenance for consistent generation efficiency. One option, explored recently, is artificial intelligence (AI) to replace conventional maintenance strategies. The growing importance of AI in various real-life applications, especially in solar PV applications, cannot be over-emphasized. This study presents an extensive review of AI-based methods for fault detection and diagnosis in PV systems. It explores various fault types that are common in PV systems and various AI-based fault detection and diagnosis techniques proposed in the literature. Of note, there are currently fewer literatures in this area of PV application as compared to the other areas. This is due to the fact that the topic has just recently been explored, as evident in the oldest paper we could obtain, which dates back to only about 15 years. Furthermore, the study outlines the role of AI in PV operation and maintenance, and the main contributions of the reviewed literatures.



Citation: Abubakar, A.; Almeida, C.F.M.; Gemignani, M. Review of Artificial Intelligence-Based Failure Detection and Diagnosis Methods for Solar Photovoltaic Systems. *Machines* **2021**, *9*, 328. <https://doi.org/10.3390/machines9120328>

Academic Editor: Christoph M. Hackl

Received: 26 October 2021

Accepted: 26 November 2021

Published: 1 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: artificial intelligence; photovoltaics; fault detection; machine learning; operation and maintenance; renewable energy

1. Introduction

The rapid development of technology and social advancements has led to the skyrocketing of energy demand, which has, in turn, resulted in an increase in fossil fuel generation of energy [1,2]. This has raised concerns of high CO₂ emission into the atmosphere due to the combustion of fossil fuels [3,4], which leads to global warming, GHG emissions, climate change, and other environmental issues [5]. Owing to the global commitment to overcome these issues by reducing fossil fuel energy generation to the bare minimum, the renewable industry has experienced an exponential growth and development in recent years. Renewable energy sources, especially solar [6,7], have been increasingly adopted for residential, commercial, and industrial applications [8–10]. The 2020 first quarter (Q1 2020) report of the National Renewable Energy Laboratory (NREL) stated that at the end of 2019, the installed solar PV capacity totaled 627 GW_{DC}, an increase of 115 GW_{DC} from the previous year [11].

Solar PV systems, however, need constant maintenance in order to efficiently operate over time. Therefore, strategies have to be in place to effectively monitor and maintain these systems. Various conventional methods are deployed by experts to carry out preventive, corrective, and predictive maintenance activities [12]. These methods usually equip the PV system with ground fault detection interrupters (GFDI) and overcurrent protection devices (OCPD). However, most of the time, they are not sufficient enough for detecting certain faults due to low irradiance conditions, nonlinear output characteristics, PV inverter maximum power point trackers (MPPT) or high fault impedances [13]. The need for more adequate and intelligent strategies of detecting and diagnosing faults in PV systems has

encouraged the adoption of AI-based methods. These methods utilize machine learning to train models to detect and locate various faults, monitor the general health status, and help the maintenance engineers of PV systems to rapidly expedite system recovery [13].

Over the years, AI has been utilized in PV system's design [14], MPPT [15], power prediction [16], as well as failure detection and protection [17–21]. The authors of [22] detailed the key characteristics that distinguish fault detection and diagnosis methods in PV systems. According to the authors, these characteristics include rapidity in detecting defects, input data (climate and electrical data), and selectivity (ability to distinguish between different faults). In addition, the study divided AI-based fault detection and diagnosis methods into two categories. First, visual and thermal methods for detecting discoloration, browning, surface soiling, hotspot, breaking, and delamination. Second, electrical methods for detecting and diagnosing faulty PV modules, strings, and arrays, such as arc faults, grounding faults, and diode faults. In [17], the adaptive neuro-fuzzy inference framework was adopted for the development of a smart fault detection approach for PV modules, while in [18], a dual-channel convolutional neural network model with a feature selection structure was proposed for PV array fault diagnosis. K-NN was used by the authors of [19] for the modeling of PV systems based on experimental data for detection, while in [20], a wavelet-based protection strategy was presented for the detection of a series of arc faults interfered by multicomponent noise signals in grid-tied PV systems. In [21], a PV fault detection algorithm that integrates two bi-directional input parameters based on the artificial neural network was presented. A novel extreme learning machine (ELM)-based modeling method, featuring high training speed and generalization performance was proposed in [23], using current-voltage (I-V) curves measured at different operating conditions, for the characterization of the electrical behavior of PV modules. The authors in [24] attempted to improve the integration of PV systems into the electrical network by controlling the converter and inverter. This is achieved through the introduction of an adaptive reference PI (ARPI) for the inverter aimed at enhancing the system performance by supporting low voltage ride through (LVRT) capability and smoothing of the PV power fluctuations during variable environmental conditions. The authors of [25] proposed a model based on the geographic information system (GIS)-based reinforcement learning, for the optimal planning of rooftop PV system. The model considers the uncertainty of future scenarios across the buildings lifecycle. Another AI application is PV systems, which is utilized in [26] for the optimal dispatch of PV inverters in unbalanced distribution systems using reinforcement learning, while the authors of [27] used AI for the optimal design of a phase change material integrated renewable system with on-site PV, radiative cooling, and hybrid ventilations. In addition, reinforcement learning with the fuzzified reward approach was used in [28] for controlling the MPPT of PV systems. Another MPPT algorithm of the PV system, which is based on irradiance estimation and the multi-Kernel extreme learning machine, was presented in [29] in order to reduce investment costs and improve PV system efficiency, while in [30], the deep reinforcement learning approach was used for MPPT control of partially shaded PV systems in smart grids. Two other literatures on the application of AI for MPPT control are provided in [31,32]. The authors of [31] presented a new combined ELM variable steepest gradient ascent MPPT for the PV system, while the authors of [32] presented a novel meta-heuristic optimization algorithm based MPPT control technique for partially shaded PV systems. As stated earlier, AI was also utilized for PV power prediction as in the case of [33], where the short-term PV power prediction was achieved using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets. Moreover, AI was used for a similar application in [34], where the deep learning and wavelet transform integrated approach was used for short-term solar PV power prediction. The authors of [35] developed an intelligent real-time power management system, where an incremental unsupervised neural network algorithm was used to predict the output power and then detect the power fluctuations occurrence of a grid-tied PV system. A comparative study on short-term PV power prediction using the decomposition based ELM algorithm was presented in [36],

while in [37], a short-term PV power forecasting model using the hybrid deep learning approach was proposed.

This study presents an extensive review of AI-based methods and techniques of fault detection and diagnosis reported in various literatures. The contribution of the study is in outlining the characteristics of the reviewed AI-based methods and their effectiveness in rapidly and efficiently detecting faults with minimal error, since the effectiveness of a fault detection and diagnosis method depends on the following factors: Its ability to detect a fault and pinpoint its location in the shortest possible time; its relative affordability; and ease of use [38]. The structure of the remaining part of the paper is as follows. Section 2 discusses the various types of faults that occur in PV systems; Section 3 introduces artificial intelligence and machine learning; Section 4 provides a review of the AI-based fault detection and diagnosis methods proposed in various literatures; and Section 5 concludes the present work and discusses its perspective.

2. Types of PV System Faults

Over time, PV systems experience fault occurrences that affect the system's operating efficiency, may cause damage to the system components, and may also lead to dangerous fire threats and safety hazards. PV system faults are classified as physical, environmental or electrical faults [39]. Panel faults, such as PV cell internal damages, cracks in panels, bypass diodes, degradation faults, and broken panels are classified as physical faults [39]. Shade faults due to bird dropping, dust accumulation, cloud movement, and tree shadows are classified as environmental faults [39]. Faults that are classified as electrical faults include MPPT faults, open-circuit faults, ground faults, line-line faults, short-circuit faults, arc faults, and islanding operation [39,40]. This section briefly discusses the different types of faults peculiar to PV systems.

2.1. Shading Faults

Shading occurs when objects, such as trees, neighboring buildings, and overhead power lines, cast shadows on PV modules [41,42]. Shading in PV arrays could be homogeneous, where there is a balanced reduced irradiation across the PV panels or non-homogeneous, where there is an unbalanced reduced irradiation across the panels [39].

2.2. Arc Faults

A frequent high-power discharge of electricity through an air gap between conductors causes this type of fault [43–45]. The two forms of arc faults include first, the series arc faults that usually originate from solder separation, connection corrosion, cell damage, rodent damage or abrasion from numerous sources. Second, parallel arc faults that result from insulation failure in current-carrying conductors [43,44].

2.3. Line-Line Faults

A line-line fault is an unintentional short-circuit between two points with differing voltage potentials [46–48]. These faults are more difficult to detect than other faults and are frequently misinterpreted as short-circuit faults in grounded PV systems, since the fault current is determined by the voltage differential between two fault spots [39]. The most common types of line-line faults are intra string faults, which are short-circuit faults between two locations on the same string, and cross string faults, which are short-circuit faults between two places on separate threads [39].

2.4. Ground Faults

To protect users from a possible electric shock, it is common practice that the metallic parts of the PV array are grounded using earth-grounding conductors (EGC) [39]. The term "ground fault" refers to any unintentional connection between a current-carrying conductor and an EGC that results in a current flow to the ground [45,46,49,50].

2.5. Other Faults

This refers to any fault that cannot be categorized under any of the previously discussed faults. These types of faults include MPPT and inverter faults that mostly occur due to inverter components failure, such as IGBTs, capacitors, and converter switch failure [51–53]; bypass diode faults resulting from a massive reverse current flow during faults, which leads to short-circuits [54,55]; blocking diode faults, also as a result of a reverse current flow [39]; open-circuit faults caused by items falling on PV panels, physical failure of panel-panel cables or joints, and sloppy termination of cables, plugging, and unplugging connectors at junction boxes [56]; faulty connections damage of connecting cables or a wrong connection of panels [57]; battery bank failures due to abnormal charging conditions [39]; and blackouts caused by natural disasters, such as a storm and lightning [58]. Most of these mentioned faults are usually due to an after effect of the other faults [39].

3. Artificial Intelligence and Machine Learning

Machine learning (ML) is a subsection of AI, which, in turn, is one of the most recent branches of science and engineering that emerged shortly after the Second World War, in order to attempt to understand and build intelligent entities [59]. ML thrives on extracting meaning from a large amount of data. Therefore, ML refers to the tools and technology that can be used to answer questions using the currently available data. Philosophy, statistics, biology, computational complexity, information theory, artificial intelligence, cognitive science, and control theory are some of the fields that machine learning draws on concepts and results [60]. In recent years, numerous machine learning algorithms and applications have been successfully created and utilized for different applications from autonomous vehicles, where these algorithms learn to drive on public highways, detect fraudulent transactions using data-mining programs, and are applied to information-filtering systems. This has led to critical advancements in the foundations of this field, which are theory and algorithms.

According to [61–63], machine learning is a branch of artificial intelligence (AI) that allows a system to learn from experience and improve without the need to be explicitly programmed. Its goal is to create computer programs that can access data and learn on their own [61]. Basically, the concept of ML is the use of fed-in data and algorithm to generate artificial knowledge, which is guided by sets of pre-defined analytical rules for pattern recognition in collected data [60]. In simple clear terms, ML is the use of quantitative and qualitative data to answer questions with ease and precision. Using ML, computers and IT systems are able to perform tasks, such as discovering, extracting, and summarizing relevant data, making predictions based on analyzed data, independently adapting to certain developments, calculating probabilities for specific results, and optimizing processes based on recognized patterns [61].

AI and ML are found in our everyday activities from transportation in the form of self-driving and self-parking cars, image recognition, google searches, fraud detection in the banking and finance sector, diagnosis and prescription in medical fields, recommendation systems, as well as social media applications [59]. In PV system applications, ML and AI are used in predicting solar radiation, sunshine duration, and clearness index; mean temperature; and insolation and diffusion fractions [64]. In addition, they are used for the sizing, configuring, modeling, simulation, and control of PV systems, detection and diagnosis of faults, and the forecasting of the output electricity from standalone and grid-connected PV systems [22,64–72].

3.1. Machine Learning Types

3.1.1. Supervised Learning

In this ML type, models are defined in advance and systems learn from the given input and output pairs, i.e., the input data and desired output are labeled [61]. With enough data knowledge, one can help the machine connect the dots with supervised learning using the labeled sample data and correct output.

3.1.2. Unsupervised Learning

Here, the AI learns without the aid of predefined target values, i.e., the model is required to identify patterns in an unlabeled input data [61]. Learning and improving by trial and error is key to unsupervised learning. Unlike supervised learning, here you are not working with labeled data, you are not showing the machine the correct output. You are using different algorithms to let the machine connect the dots by studying and observing data. In unsupervised learning, the chances of the machine to find patterns or classifications that humans can never see is very high.

3.1.3. Semi-Supervised Learning

This is a combination of supervised and unsupervised learning advantages [61]. Here, training starts with a small amount of dataset in order to allow the machine to get familiarized with the data. In addition, the machine studies and observes the data to expand its vocabulary/database using inductive reasoning. Another form of semi-supervised learning is the transductive reasoning, which allows one to narrow down the unlabeled data using unknown knowledge of collected data. Semi-supervised learning is not very common in machine learning applications.

3.1.4. Reinforcement Learning

In this situation, the model is granted autonomy to engage with a dynamic environment that gives feedback based on rewards and punishments, i.e., the model is taught through positive and negative interactions [61]. This method of learning differs significantly from the other three methods. The machine iterates until the outcome is enhanced each time, coming closer and closer to high-quality output.

3.1.5. Multitask Learning

Multitask learning helps several algorithms share their experience with each other, thereby helping them learn concurrently rather than individually [73].

3.1.6. Ensemble Learning

Ensemble learning is a combination of two or more algorithms that form one single algorithm [74]. Here, it has been observed that a collection of algorithms almost always outperforms an individual algorithm when carrying out a particular task [75].

3.1.7. Neural Network Learning

Neural networks, also referred to as artificial neural networks (ANN), are derived from the biological concept of brain cells called neurons. Therefore, to understand ANN, one has to be familiar with how neurons work [74]. An ANN functions on three layers (input layer, hidden layer, and output layer), in the same way as the brain neurons work on four parts (dendrites, nucleus, soma, and axon) [74]. The input layer receives the data, which is then processed by the hidden layer, before it is sent as a calculated output to the output layer [76].

3.1.8. Instance-Based Learning

In this case, the algorithm learns a specific pattern and then applies it to new data [74]. This learning method becomes more sophisticated as the amount of data grows [77].

3.1.9. Evolutionary Computation

Evolutionary computation is currently a distinct branch of artificial intelligence inspired by nature [78,79], with smart methods based on evolutionary algorithms targeted at solving various real life problems through natural processes involving live things [79]. It is based on random processes, data regeneration, and data replacement within a system, such as a personal computer or any other data center. A variety of evolutionary computa-

tion approaches are utilized for different applications, including image processing, cloud computing, and grid computing [79].

4. Artificial Intelligence-Based Failure Detection and Diagnosis Methods

Most of the AI-based methods used in failure detection and diagnosis adopt ML models, such as support vector machines, wavelets, neural networks, fuzzy logics, decision trees, graph-based semi-supervised learning (GBSSL), regression, etc. for the development of models and algorithms that are trained to learn the relationships between the input and output parameters of PV systems. The data are obtained from experimentally accurate PV model measurements, and then split into training and test datasets. This section reviews and discusses the AI-based detection and diagnosis techniques proposed in various literatures. Figure 1 shows a distribution of the available literature on fault detection and diagnosis methods for PV systems based on the mentioned ML models. It presents a rough estimate of the number of conference and journal papers published to date. In general, the amount of publications on the subject is relatively low, which is due to the fact that it is a newly explored area of research that dates back to only about 15 years.

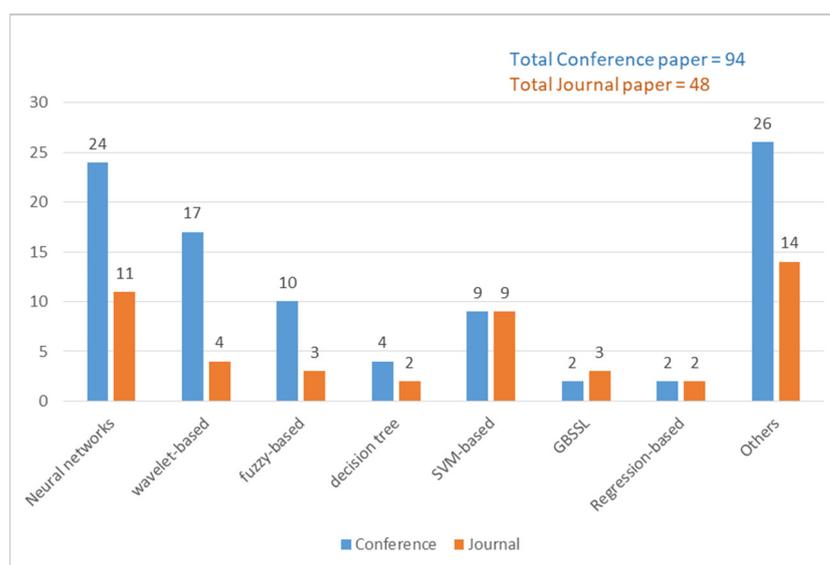


Figure 1. Distribution of AI-based fault detection and diagnosis methods for PV systems found in the available literature.

4.1. Neural Network-Based Methods

The study presents a literature review concerning the AI methods of fault detection and diagnosis that are neural network based. Table 1 presents a summary of the different methods presented in the subsection, pointing out the main contribution of each reviewed literature.

In [80], a simple and effective Bayesian neural network (BNN) model for estimating power losses in PV plants owing to soiling was devised. Four models were built based on the Bayesian neural network (BNN) to assess the performance of two plants for dirty and clean module conditions under standard test conditions (STCs). The loss due to the soiling impact is shown by the difference in the STC power between the two conditions. The study found that utilizing a BNN model rather than a polynomial model for calculating the STC power of a PV system is more successful due to various factors that affect the polynomial model's performance, including the database size. Figure 2 presents a schematic diagram of the BNNs used for calculating the STC power, consisting of an input layer (which has solar irradiance and cell temperature as input), a single hidden layer (estimated during the training process), and an output layer (which provides the STC power output produced by the plant). Bayesian regularization, a process of updating the weight and bias values

according to the Levenberg-Marquardt optimization technique, which helps in reducing a combination of squared errors and weights, and determining the correct combination to produce a network that generalizes well, can greatly improve the generalization ability of neural networks [81]. BNN is basically back propagation (BP) with an additional ridge parameter added to the objective function [80]. The study provides an important contribution as it helps the operation and maintenance personnel in decision making between washing cost and losses in energy production.

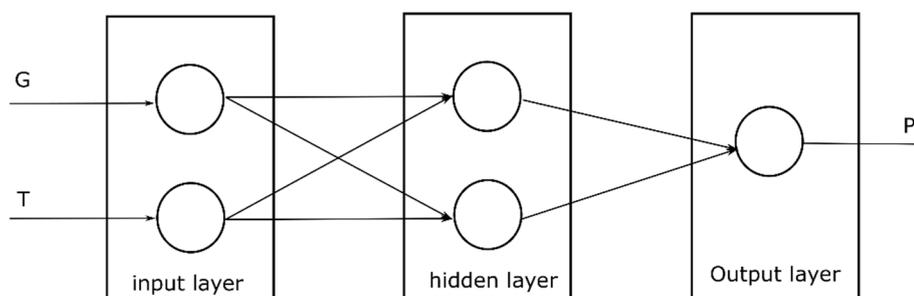


Figure 2. Schematic block diagram of Bayesian regularization [80].

In order to successfully detect and categorize PV array faults, the authors of [82] employed a deep two-dimensional convolutional neural network (CNN) to extract features from two-dimensional scalograms generated from the PV system data. The study took into account five different fault scenarios as well as the use of MPPT. There are two variations of the proposed method. First, the last layers of a pre-trained AlexNet CNN [83] are fine-tuned to generate a six-way classifier in the first configuration. In the second configuration, features are extracted from a specific layer of a pre-trained AlexNet and then combined with a classical classifier. The suggested model's performance is compared to machine learning- and deep learning-based models. The suggested method surpasses the previous methods in terms of detection accuracies for both noisy and noiseless data, according to the authors. They also illustrated the need of representative and discriminative features for categorizing errors (rather than using raw data), especially in noisy environments. Automatic feature extraction based on deep learning has been found to be superior to manual feature extraction. In order to better explain the proposed method, Figure 3 presents a flowchart showing the proposed PV array fault diagnosis method and existing methods. Another neural network-based method of fault detection and diagnosis in solar PV systems, which uses Elman neural network (ENN), is presented in [84]. The study examines the implicit mining link between original data and fault types, develops multiple hypothesis models, and analyzes the mean and variance of diagnostic errors to determine which diagnostic model is optimal. The suggested fault diagnosis approach based on ENN overcomes the problem of PV system multi-source and multi-type defect identification by minimizing the number of sensors, which only collects PV operation data and data from the atmospheric environment.

The main contribution of [66] is the proposition of a technique for isolating and identifying faults that occur in the PV system, and its implementation into a field programmable gate array (FPGA) with real-life application effects. The proposed approach detects and diagnoses faults that occur in PV bypass diodes, cells, modules, and strings. It accomplishes this by examining a set of parameters, such as current, voltage, and the number of peaks in the I-V characteristics that indicate normal and abnormal PV system operation. Two separate algorithms are used in this strategy. The first algorithm isolates defects with different combination attributes using a signal threshold approach. The second technique uses an ANN-based approach to identify errors that have the same mix of features. The approach is low-cost and easily adaptable to large-scale PV systems. The block diagram of the proposed fault detection technique based on the threshold approach and ANN is presented in Figure 4.

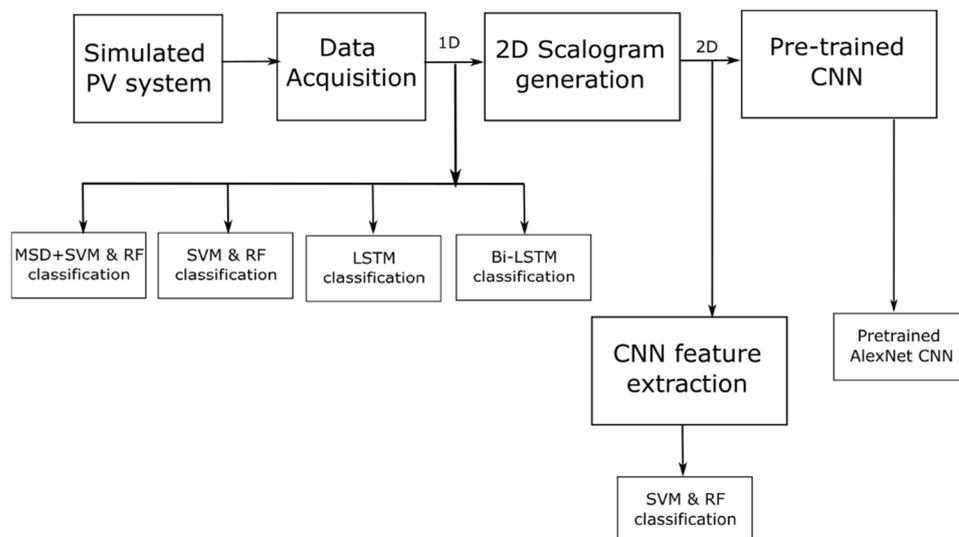


Figure 3. Flowchart showing the proposed method of PV array fault diagnosis and existing methods [82].

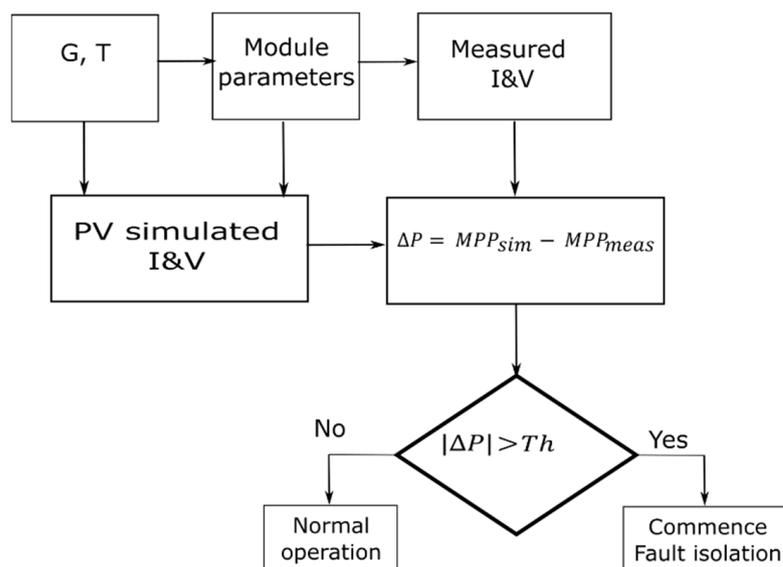


Figure 4. Block diagram of the proposed fault detection technique based on the threshold approach and ANN [66].

A method for detecting the islanding phenomenon in the PV system was introduced in [85]. To detect islanding actions, the method employs a multi-variable method based on an extended neural network that combines passive and active detection modes. The method combines the extension theory’s extension distance with a neural network’s learning, recalling, generalization, and parallel computing capabilities. The study used an extension neural network to distinguish between power quality interference (voltage swells, voltage dips, power harmonics, and voltage flickers) and actual islanding operations at the grid power end, in order for the islanding phenomenon detection system to cut off the load correctly and promptly when a real islanding operation occurs. The detection algorithm is created and translated using a PSIM software package based on the C language and written in dynamic-link library (DLL) modules. The signals sent by DLL are passed back to the controller to complete the islanding detection control. An enhanced machine learning based approach for the detection and diagnosis of short-circuit faults, and a complete disconnection of the string from an array in the DC side of the grid-connected PV system is

presented in [86]. The process uses a probabilistic neural network (PNN) classifier with one diode model (ODM) and a parameter extraction method to create a trustworthy model of a real-world PV system. Two PNN classifiers are used in the proposed method, one is for detecting fault occurrences and the other for diagnosing the type of fault. There are four stages to the method: Array parameter extraction; experimental model validation; elaboration of database of both healthy and problematic operations; and network design, training, and testing based on the best-so-far ABC algorithm. The contribution of this study is highlighted in the model's ability to detect a fault, while also pinpointing its origin. However, for the method to be effective, the high-quality database which is not always readily available, is required to deal with classification problems. To deal with this issue, the authors suggested having in place a trusted simulation model, which is able to mimic the exact healthy and faulty behaviors of a PV system. Another PNN-based intelligent method for PV system health monitoring was proposed in [56], which can detect and categorize short- and open-circuit faults in real time, as well as locate the faulted PV string in a grid-connected PV system. To detect and diagnose faults, the suggested technique uses data obtained from various sensors in PV systems, such as voltage, current, irradiation, and temperature which are used to deliver information on fault occurrence. The PNN used in the method has four layers: The input layer (the number of neurons in the layer represents the number of training and testing samples); the hidden layer (whose pattern units are equal to the training set sample space); the summation layer (the number of neurons is equal to the number of sample space classes); and the output layer or the decision layer (containing one neuron which provides the classification decision). Moreover, it was developed and validated in computer programs utilizing a novel approach to PV system modeling that only requires data from the manufacturer's datasheets provided under normal operating cell temperature conditions (NOCT) and STCs. The modeling approach is an improvement to the previous approaches where STC conditions, I-V characteristics or NOCT conditions are used but never combined together. This systematically builds a relationship between the ideality factor, thermal voltage, and series resistance with the PV module temperature using the manufacturers' datasheet elements. The PV system simulation model is then used to implement and validate the PNN-based detection model and classification method. The authors of [87] explored real time online fault detection for PV modules under partial shading conditions. The approach suggested in the paper is an intelligent method that uses artificial neural networks (ANN) to estimate the output PV current and voltage under varying operating conditions utilizing solar irradiance and cell temperature meteorological factors. Since it performs the real time correlation of estimated performances with the measured performances under variable conditions, the method can also be used to detect the possible anomalies in PV modules. The model proposed is independent of the measured PV module performance which makes its system of fault detection autonomous. Figure 5 depicts the fault detection flowchart of the proposed method. The results of the study show that the proposed method can accurately estimate the output and detect any decrease in output power without requiring any complex calculations or mathematical models. However, it does necessitate that the ANN be trained on a regular basis in order to accurately estimate the output parameters. The approach could also be used in PV arrays or large-scale PV plants, as well as in low-cost microcontrollers for real-time applications.

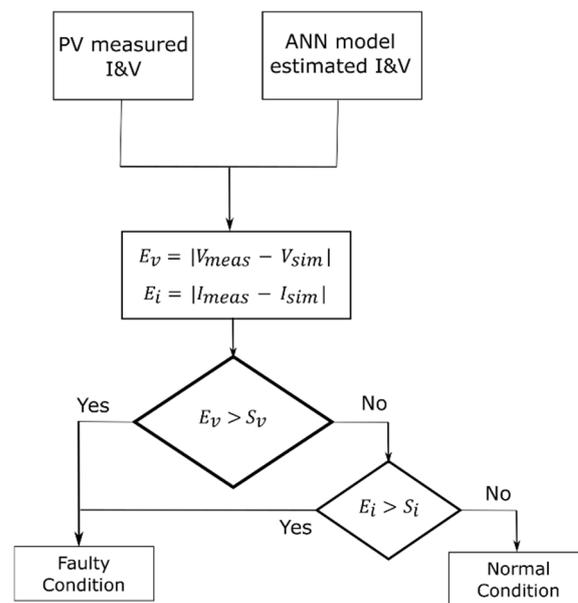


Figure 5. Proposed fault detection flowchart diagram [87].

According to the authors in [88], few data collecting systems in solar power plants focus on intelligent reasoning of the plant's state, despite the fact that the data of these collecting systems offer a wide range of capabilities. To tackle this problem, the authors presented a description of a novel data acquisition system. A Bayesian belief network-based (BNN) fault detection and diagnostic system was then built, which analyzes the acquired data for the existence of faults and intelligent reasons for potential causes of the detected faults. The BNN-based model uses a graphical representation of a problem in the form of a hierarchical network, with nodes representing random variables and directed arcs expressing the probability regarding the dependencies between these variables. Each node has a set of states, each with a probability distribution associated with it. Each arc reflects a conditional probability based on the preceding nodes. A node's state could be as simple as true or false in the most basic scenario. In a more complicated instance, the set of states could include multiple discrete states, such as low, medium, high, and very high. The measurements obtained by a BBN-based inference engine produce a change in the probability values in the respective nodes, which impacts the connected nodes in the network and leads to the automatic derivation of a decision on the likely reason of a failure. By developing a framework for analyzing sensor results and translating them to a Bayesian belief network using the Netica API, the system lays the groundwork for future advancements. The Fault Identification for Nasa Exploration Missions and Navigation (FINEMAN) system was developed as an integrated package with four main components: A connectivity interface for remotely retrieving data from the plant's data acquisition system; a preprocessor for relevant measurement selection; a fault injector for failure testing simulation; and a Netica implemented BBN interface engine. The research adopted in [64] was also ANN-based, and unlike in [87], where ANN was used for the PV module under partial shading conditions, the approach in [64] utilized AI technology for automatic MPPT fault detection and failure type judgement. The method requires five features (solar irradiation, installed capacity, MPPT power, MPPT voltage, and MPPT current) for machine learning. Each inverter's MPPT performance is gathered and modeled using a machine learning algorithm rather than a rule-based programming approach. In addition, an inefficient MPPT is discovered by comparing the real and expected power output. Moreover, faulty equipment is automatically recognized, a knowledge-based system determines the type of failure, and an alarm with the failure diagnostic is communicated to the user by a mobile device or email. The field sites collect data on power production every 5 min. As a result, the entire computation and communication process in the system takes only 5 min to complete. Using this AI

method, the authors were able to efficiently manage hundreds of projects at the same time, while also optimizing O&M performance with minimal work and resources. A novel algorithm utilizing the genetic algorithm to optimize the topology of ANN was introduced in [72]. The new algorithm is an online optimized neural network-based fault diagnostic and repairing system, aimed at providing a solution to the problem of complexity and high costs, associated with the other fault detection methods. The method offers the following factors: High speed diagnostic process as it diagnoses multiple faults in parallel; can remotely replace faulty components with good ones; can be used for modern complex PV systems; and has the ability to partition the PV panel into sub-areas. As a result, the diagnostic procedure is divided into two parts. The first is concerned with identifying the failure where the proposed method diagnoses a PV system using a genetic algorithm to optimize the topology structure of the neural network. The system is implemented in five steps: (1) Chromosome representation, which uses binary digits for topology network representation; (2) initial population, which is constructed from random individual sizes; (3) cross operator, which combines the parents to obtain two offspring using the uniform crossover; (4) mutation operation, where random mutation is used for genetic algorithm; and (5) fitness function, which is used to minimize the error value of the neural network representing the fitness function of the genetic algorithm. The second part of the proposed method is solely concerned with determining the cause of the failed area by dividing the structure of the PV system into three modules (the PV panel module, the charger module, and the battery module). Following this step, the ANN is trained with multiple types of faults for the three divided PV system structures and then the ANN begins the diagnostic process. When the proposed technique's results were compared to fuzzy-based and classic neural network-based diagnostic systems, it was discovered that the proposed method produced better results. The authors of [89] presented a neural network-based method for modeling the relationship (MPP) of a shaded PV array and environmental parameters (solar irradiance, sun angle, ambient temperature) in non-uniform settings. Similar to the shading factor, this neural-network-based function can characterize the shadow impacts on a solar PV array over time. As a result, the neural network model is able to eliminate the inaccuracy produced by the shading factor's complex calculation. In contrast to the prior efforts that only address the uniform shadow on the PV array, this method considers the non-uniform shadow and illumination. The proposed method's procedure is as follows:

- The shadow ratio is defined by the solar height and solar azimuth angles, which may be simply calculated from the time of day for a specific geographic location. Therefore, the neural network's inputs are the sun's irradiation levels, angle, and ambient temperature. The neural network's output has the highest solar PV array output power.
- Experimental data are acquired by taking measurements numerous times a day, for several days, while the solar PV array is partially shaded by a nearby object. One set of measured data is used to train the neural network, while the other is used to test it. The neural network's accuracy in estimating the PV array's maximum output power is tested using the test data.
- With low computational effort, the neural network can forecast the output power of solar PV arrays at any solar irradiation level, at any time of day, and at varied ambient temperatures over a long period of time.

The authors of [90] proposed a PV prognostics and health management (PHM) technique. The system was created to track the health of photovoltaic systems, measure degeneration, and provide maintenance recommendations. It employs a system-specific ANN model, which eliminates the need for prior knowledge of system components and design. Two detection techniques were tested in order to better monitor the health of the PV system. The energy loss fault detection system uses a neural network model to compare the sum of power loss over a lengthy period of time. An alarm threshold can be set to detect the long-term effects, such as soiling or material degradation, and alert the user to the need for maintenance. In the event of a catastrophic system failure, such as the loss of a

string of modules or an inverter failure, the acute fault detection technique examines the potential of the PV system performing below model expectations and should warn the user. When the two techniques are combined, the short- and long-term PV system defects can be detected. The metrics of the two combined methods can also be used for pre-emptive inspection and maintenance, as they allow the system operator to identify the PV system failure precursors linked to failure modes.

Table 1. Summary of neural network-based methods.

Authors	Reference	Year	Contribution
A.M. Pavan, A. Mellit, D. De Pieri, S.A. Kalogitou	[80]	2013	Evaluation of the effect of soiling in large PV plants using BNN and regression polynomial
F. Aziz, A. Ul Haq, S. Ahmad, Y. Mahmoud, M. Jalal, U. Ali	[82]	2020	Classification of line-line fault, open-circuit fault, partial shading, and arc fault using the convolutional neural net
G. Liu, W. Yu	[84]	2018	Diagnosis of open-circuit fault, overall shading effect, and partial shading effect using the Elman neural net
W. Chine, A. Mellit, V. Lugh, A. Malek, G. Sulligoi, A.M Pavan	[66]	2016	Identification of eight different faults using ANN
K.H. Chao, C.L Chiu, C.J. Li, Y.C. Chang	[85]	2011	Islanding phenomenon detection in PV systems using ANN
E. Garoudja, A. Chouder, K. Kara, S. Silvestre	[86]	2017	Detection and diagnosis of faults in the DC side of PV systems using the probabilistic neural network
M.N. Akram, S. Lotfifard	[56]	2015	Modeling and health monitoring of the DC side of PV systems using the probabilistic neural network
H. Mekki, A. Mellit, H. Salhi	[87]	2016	ANN based modeling and fault detection of partially shaded PV modules
A. Coleman, J. Zalewski	[88]	2011	Intelligent fault detection and diagnosis in the PV system using the Bayesian belief network
M. Chang, C. Hsu	[64]	2019	PV O&M optimization using ANN
A. H. Mohamed, A.M. Nassar	[72]	2015	PV system fault diagnosis using the genetic algorithm-optimized ANN
D.D. Nguyen, B. Lehman, S. Kamarthi	[89]	2009	Performance evaluation and shadow effects in PV arrays using ANN
D. Riley, J. Johnson	[90]	2012	Prognostics and health management of PV systems using ANN

4.2. Regression-Based Methods

In this subsection, PV system fault detection methods that are regression-based are reviewed. Table 2 summarizes these methods, highlighting their main contributions.

For the detection of abnormal situations in the PV system, a novel approach has been employed in [91] based on regression and SVM, to compute the ideal power generation, which takes into account all of the three categories of failures in PV systems, namely failures of PV modules, inverters, and other components. Furthermore, the proposed method makes use of variables that are already present in a small-scale PV system, eliminating the need for the installation of extra expensive sensory equipment. Power, voltage, and current are collected from the power conversion system (PCS). Solar irradiance on the surface of the PV panels is measured with a pyranometer, and ambient and PV cell temperatures are monitored with a thermometer. As a result, the suggested PV abnormal condition detection system can be efficiently used in a small-scale PV system or as an early warning system for the PV operator/owner to undertake further system inspections. Merged regression and support vector machine (SVM) models were used to create the PV abnormal condition detection system. The regression model is used to calculate the expected power generation for each solar irradiation, which is then fed into the SVM model. The SVM model, on the

other hand, uses numerous variables, including the expected power generation, which is generated from the regression model to determine the abnormal condition of the PV system. Since the data used as the model's input variable is acquired from the PCS, the proposed model does not require the installation of extra measurement devices and can be constructed at minimal cost. The detection system's accuracy is also increased by taking into account the daylight time and interactions between the independent variables, as well as using the multi-stage k-fold cross-validation technique. The proposed detection system is evaluated using real data from a PV site, and the findings show that it can successfully discern between normal and abnormal PV system conditions using basic measures. The authors in [92] proposed a condition monitoring technique based on an online regression PV array performance model, in which PV array production, POA irradiance, and module temperature/maximum power (MPP) measurements are collected during the system's initial learning phase and used to automatically parameterize the system online using regression modeling. After the model has been automatically parameterized and optimized using the regression modeling methods, the condition monitoring system enters the normal operating phase, in which the performance model is used to anticipate the PV array's power production. The authors claimed that using projected and measured PV array output power values, the condition monitoring system based on the Sandia array performance model (SAPM) [92,93], could detect power losses in the PV array of larger than 5%. When compared to the existing model-based condition monitoring systems, the suggested method is unique in that it can take advantage of the I-V scanning capabilities of a new generation of commercial PV inverters [92], according to the authors. Using the real MPP collected from the I-V curve, as well as the ambient condition sensors, the system can generate an accurate performance model of the PV array in question during field operation. The authors' proposed method has the following advantages: Simple commissioning and operation requirements; potential applicability to a wide range of PV system configurations; and it does not require modeling and testing of the PV array prior to installation. It does, however, have an initial commissioning phase where ambient conditions and array MPP measurements are used for the automatic commissioning of the PV array. Figure 6 presents the condition monitoring systems' learning or commissioning phase.

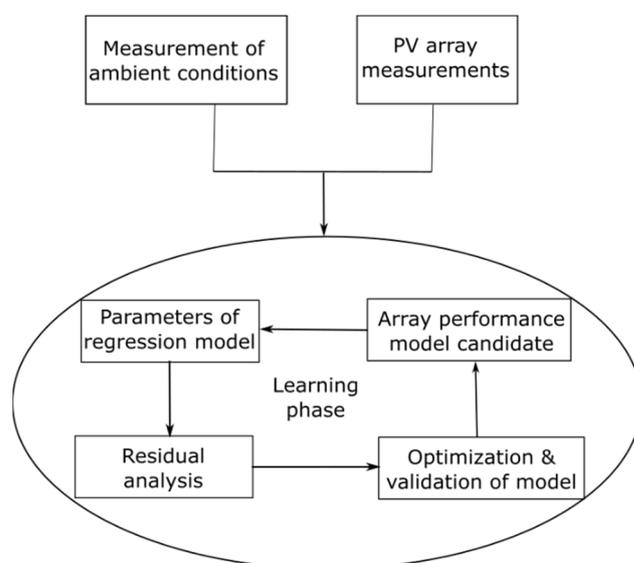


Figure 6. Condition monitoring systems' learning phase [92].

Unlike Reference [92], which only used a regression model, Reference [94] used a combination of linear regression and artificial neural networks, as well as solar irradiation, ambient temperature, and maximum power point (MPP) characteristic variables of PV modules obtained from I-V tracers at the PV installation site, to predict the performance

of soiled PV modules. In the study, two methods (linear regression and neural networks) were used in modeling the output of soiled PV panels. In the first method, multiple linear models were created based on the last cleaning cycle time stamp, thus predicting the panel's output as a function of solar irradiance. In the second method, the neural network, date, time, and irradiance, and sometimes, the temperature data were utilized and information was inputted to predict the output of the solar panel. In both methods, PV panel generation predictions were used to have high accuracies. Another regression based fault detection model is presented in [95]. The study proposed a smart algorithm for the diagnosis and prognosis of reversed polarity fault in PV generators, utilizing a hybrid optimization of a support vector regression (SVR) technique by a k-NN regression tool (k-NNR). The main contribution of the study included the development of a smart prognosis algorithm that detects, locates, and characterizes a reversed polarity fault at the cell, bypass, and string levels. For data processing in a linear space, the SVR requires a kernel function in order for the data to be transformed in a new greater dimension space. To overcome the SVR output indetermination, the study utilizes k-NNR for predicting the approximate value of the SVR undetermined output. The method was validated using 50 samples of a typical generator and the results showed homogeneous and reliable predictions.

Table 2. Summary of regression-based methods.

Authors	Reference	Year	Contribution
F.H. Jufri, S. Oh, J. Jung	[91]	2019	Detection of abnormal conditions in PV systems using the combined regression and support vector machine
S. Spataru, D. Sera, T. Kerekes, R. Teodorescu	[92]	2013	Condition monitoring of PV array based on the online regression model
S. Shapsough, R. Dhaouadi, I. Zualkernan	[94]	2018	Predicting the performance of soiled PV modules using the linear regression and BP neural network
W. Rezgui, N.K. Mouss, M.D. Mouss, M. Benbouzhid	[95]	2014	Reversed polarity fault prognosis in PV generators using the regression algorithm

4.3. Decision Tree-Based Methods

The study also presents a review of the AI methods of fault detection and diagnosis that are decision tree-based. Table 3 shows a summary of the methods presented in the subsection, with an outline of the main contribution of each literature.

A number of fault detection and diagnosis methods that are decision tree-based are provided in [67,96,97]. In [67], failure detection routines (FDRs), that use obtained datasets of grid-tied PV systems to accurately obtain and classify exhibited faults, were created. The FDRs are made of two stages, namely the fault detection stage, which uses a comparative algorithm for the detection of anomalies between the measured and simulated electrical measurements of a PV system; and a statistical algorithm for the identification of outliers, discrepancies, and normal system operation limits. The failure classification stage is otherwise called the decision stage, where developed logic and decision trees are used to perform the classification process. The failure detection stage included a comparing algorithm that discovered disparities between the measured and simulated electrical measurements on the dc side, which were obtained using empirical parametric models for each point. Significant variations between the measured and simulated parameters, indicating a noteworthy performance gap, were classified as failures. The obtained incident global irradiance at the POA, PV system specs, and module temperature were required as inputs for the detection algorithm utilized in this study. Another comparative method utilized in the study is three-sigma limit method. When typical operation limitations set by defined criteria were exceeded, failures were recognized using this approach. The upper

and lower control limits in statistical quality control charts were set using three-sigma limits (a statistical calculation that refers to data within three standard deviations of the average), which are commonly used to set the upper and lower control limits in statistical quality control charts. To establish and display the researched operation and boundaries, control charts were used. The typical operation limitations of the PV monitoring system were calculated by dividing the measured and simulated electrical data by the ratio. According to the PV production model, the ratio was used to determine how close the measurements were to their calculated values. The closer the ratio was to 1, the closer the measured parameters were to the modelled values. Since the system performance is affected by sunlight levels, an extra step is performed. To decrease the high bias errors occurring at low irradiance levels, the datasets were filtered at global irradiance levels $>50 \text{ W/m}^2$. In the statistical failure detection approach, the local outlier factor (LOF) algorithm was used to find the density-based local outliers. With LOF, each point's local density is compared to its k -nearest neighbors (k -NN), and if the point's density is significantly lower than its neighbors, then the point is in a sparser region than its neighbors, indicating that it is an outlier. Moreover, outlier testing was done using the Bonferroni outlier test algorithm. Based on the linear regression model of the observed and simulated dc power of the array, this program returns the Bonferroni p -value for the most extreme observation. In addition, the seasonal hybrid extreme studentized deviates (S-H-ESD) algorithm was used to find anomalies in a time series dataset that follows an approximately normal distribution. The S-H-ESD technique is an extension of the generalized extreme studentized deviates (ESD) algorithm that uses the time series decomposition and robust statistical measures in conjunction with ESD to detect both global and local anomalies [98]. For the detected failure classification, logic and decision trees were developed. In addition, the decision trees were trained using continuous samples divided in a 70:30% train to test the set ratio utilizing the acquired datasets that included the feature patterns seen during normal and faulty operations. Moreover, they were produced using supervised learning procedures. The accuracy of the proposed FDRs for fault detection and classification was demonstrated in the obtained results for three fault types (inverter failure, bypass diode faults, and partial shading fault), which showed that accuracy rates of 98.7, 95.3, and 96.6% were recorded for inverter failure, bypass fault, and partial shading, respectively. The authors in [96] presented a decision tree-based fault detection and classification for the PV array with an easy and straightforward model training process. Under the creation of a decision tree model for an experimental PV system in both normal and faulty working situations, the authors employed the PV array voltage, current, operating temperature, and irradiance as attributes for the training and test sets. The collected data and pre-processed training set are chosen at random from the experimental data and utilized to create the decision tree model using the WEKA software [99], after which the model is evaluated and validated on unseen real data. Fast training and classification phases, explicit interpretation, and straightforward implementation as an algorithm are all advantages of the proposed decision tree paradigm. Another benefit of the model is that it can detect problems in real time, with detection accuracy ranging from 93.56 to 99.98% and classification accuracy ranging from 85.43 to 99.8%, depending on the model's size. The authors of [97] proposed a defect detection and diagnostic technique for grid connected PV systems (GCPVS) based on the C4.5 decision tree algorithm (which is one of the most popular machine learning algorithms for classification problems [97]), in which a non-parametric model is utilized to forecast the state of the GCPVS through a learning task. Three numerical attributes (ambient temperature, irradiation, and power ratio) which are calculated from the measured and estimated power, as well as two targets (the first of which is either a healthy or a faulty state for detection, and the second of which contains four classes of labels named free fault, string fault, short-circuit fault or line-line fault for diagnosis) are chosen to form the final used data. The dataset was divided into two halves, with 66% utilized for learning and 34% for testing. Then, over the course of 5 days, additional data were collected to measure the robustness, effectiveness, and efficiency of both models. The dataset is required for the

learning process in order to construct the decision tree. As a result, an acquisition system is developed to be able to record and store data, such as climatic variation, as well as electrical variables, such as current, voltage, and power at the MPP. Three attributes are chosen, including temperature ambient, irradiation, and the power ratio, which is calculated from the estimated power by the Sandia model and the measured power of GCPVS production. The Sandia model is an empirical relationship that is used to estimate the generated power from a system in a healthy state at MPP using STC data. Since this model has unknown parameters, the flower pollination algorithm (FPA) is used to find the optimal parameter values that correspond to the smallest root mean square error between the estimated Sandia output and the measured power. As a result of the high correlation between the power ratio and the system state, a nominal property called target is constructed as a class label in each instance data in order to accurately forecast these errors. Two major approaches lead to the construction phase. To begin, a splitting criteria is used to select the best split attribute. Thereafter, the tree grows in length as this technique is repeated iteratively in order to categorize all of the instances or to verify one of the stopping criteria. Then, once the tree model has been obtained, a pruning process is carried out to remove the unneeded sub-trees in order to minimize the overfitting phenomena, which can result in a reduction in model complexity due to the reduced tree size. According to the test findings, the models have a great prediction performance in the detection with high accuracy, while the diagnostic model has an accuracy of 99.8%.

Table 3. Summary of decision tree-based methods.

Authors	Reference	Year	Contribution
A. Livera, G. Makrides, J. Sutterlueti, G.E. Georghiou	[67]	2017	Advanced failure detection algorithm and PV performance outlier decision classification
Y. Zhao, L. Yang, B. Lehman, J.F. De Palma, J. Mosesian, R. Lyons	[96]	2012	Detection and classification of PV array faults based on the decision tree algorithm
R. Benkercha, S. Moulahoum	[97]	2018	C4.5 decision tree algorithm-based fault detection and diagnosis for grid-connected PV systems

4.4. Support Vector Machine-Based Methods

The methods reviewed in this subsection are support vector machine-based. Table 4 presents a summary of these methods as well as their contributions.

The authors of [100,101] suggested a system based on support vector machines (SVM) and k-NN tools with the goal of building a fault detection and diagnosis algorithm for PV generators. The algorithms are smart, according to the authors, since they are a hybridization of the SVM approach and the k-NN tool, which are used to improve the classification rate against observations on the classifier itself. The systems' originality is in the construction of a smart classifier based on collected data from the control system, as well as the fault identification and localization of short-circuits in a cell, bypass, and blocking diodes. First, the proposed method utilizes the SVM algorithm, which is a two-class classification technique that seeks hyper-plane separating positive examples from negative ones by ensuring that there is a maximum margin between the nearest positive and negative examples. This ensures that the idea may be applied to new situations, as new examples may not be as similar to those used to determine the hyperplane, but may be on either side of the border. The selection of support vectors, which reflect the discriminate vectors by which the hyper-plane is determined, is a benefit of this method. Only those supporting vectors are utilized to assign a new case, and the examples used in the hyper-plane search are no longer required. Second, k-NN which is a simple and straightforward approach, and does not require learning is utilized to compare new examples of unknown

class to old examples in its database. Then, for this new example, k-NN chooses a majority class among its nearest neighbors. In summary, the method uses the activation function of the SVM of Gaussian type and the Euclidean distance between the gravity centers of database observations of the k-NN method. The obtained simulation results, using the proposed smart algorithm in both literatures, exhibit a high classification rate and low error rate. However, the algorithms take a longer processing period due to the mathematical computations. Therefore, future works should focus on improving this aspect of the algorithms. The authors of [48] proposed an algorithm to improve the detection accuracy of line-to-line faults in PV arrays that occur under a wide range of situations, such as low irradiance conditions, high impedance faults, and low mismatch faults. The algorithm is based on pattern recognition (multi-resolution signal decomposition (MSD)) and machine learning techniques (two-stage SVM classifier). It takes advantage of the MSD technique for the extraction of the feature space of line-to-line faults, while the SVM part is essentially for decision making. The system does not require numerous sensors, since it uses measurements of the overall voltage and current of the PV array, thus making it an economical and fast option. It detects line-to-line defects quickly and accurately, and it can be combined with fault location techniques to solve faults quickly. The MSD stage performs digital signal processing (DSP), allowing for the simultaneous time and frequency analysis of a signal, such as the analysis of both stationary and transient components of power quality disturbance. The SVM stage, on the other hand, is carried out to improve accuracy. The two-stage SVM is a binary classifier, which requires training utilizing a minimal amount of historic data from the tested PV system. The authors suggest that the proposed method is not limited to line-line faults only and could be used for the detection of other PV system faults. A method for detecting problems and monitoring the state of PV modules using a two-class data fusion method is introduced in [102]. The approach was created by combining monitoring data from sensor nodes in wireless sensor networks (WSNs) at a monitoring center with a new semi-supervised support vector machine (SVM) classifier, devised and trained using the monitoring center's existing sun irradiance big data. The monitoring center was created in order to access various monitoring data from various PV power stations, and multiple applications were created to use the envisioned system in various platforms. In this paper, a wireless monitoring subnetwork was created to retrieve crucial data from PV modules in power stations, such as the current, voltage, and temperature. The monitoring data received from each sensor was fused by a sink node with sunlight intensity information, and the fusion results were provided to the monitoring center over Internet networks. The data received from the sink node were parsed using the data access interface, and the data from the parsing process was double-checked using the outlier detection technique. The Cloud management module, which was also in charge of data security transmission between the Cloud and the applications, retained the regular data in the private Cloud. The authors built a semi-supervised SVM classifier using historical monitoring sunlight intensity data, which was employed in the outlier identification and solar power forecast algorithms. An outlier identification technique is designed using the prediction model provided by the trained classifier to identify and locate PV module faults by computing the average value of the problematic data. Furthermore, the authors employed a novel application of the PGKA technique to ensure the security of data transmission between the Cloud and apps. The fact that this approach does not require third-party certification to maintain file encryption and encryption keys is an apparent benefit. Despite the fact that this research focuses on crucial PV power station challenges, there is still a long way to go in terms of gathering PV power station data and intelligence properly.

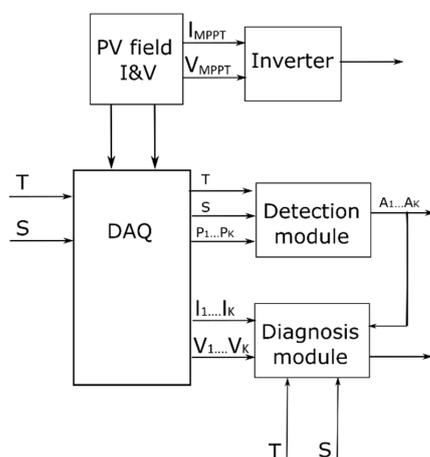
Table 4. Summary of SVM-based methods.

Authors	Reference	Year	Contribution
W. Rezgui, H. Mouss, K. Mouss, M.D. Mouss, M. Benbouzid	[100]	2014	Short-circuit fault diagnosis in PV generators using a smart SVM algorithm
W. Rezgui, N.K. Mouss, L.H. Mouss, M.D. Mouss, Y. Amirat, M. Benbouzid	[101]	2014	Short-circuit and impedance faults smart diagnosis in PV generators using the k-NN optimized SVM classifier
T. Hu, M. Zheng, J. Tan, L. Zhu, W. Miao	[102]	2015	PV system intelligent monitoring based on solar irradiance big data and wireless sensor networks

4.5. Neuro-Fuzzy-Based Methods

This subsection presents a review of AI methods of fault detection and diagnosis that are based on the neuro-fuzzy technique. Table 5 presents a summary of the methods as well as their contributions.

The authors of [103] proposed an intelligent system for automatic fault detection in PV fields based on the Takagi-Sugeno-Kahn fuzzy rule-based system (TSK-FRBS) [104]. The method is based on the analysis of recorded voltages and currents collected from a PV plant's inverter. The TSK-FRBS is a power estimator module that estimates the PV field's immediate power production in normal operating conditions (using temperature and irradiance input signals to assess the DC power that the PV plant should produce) and compares it to the real power to check for differences. If there is a large disparity between the two power sources, an alarm signal is issued. In this circumstance, the TSK-FRBS has two advantages. First, it can describe complicated system behaviors without requiring the use of a mathematical model. Second, it is able to deal with noisy and vague data. The schematics of the proposed intelligent system, which is connected to a PV system is shown in Figure 7. It consists of the data acquisition module, the detection module, and the diagnosis module, in a multi-array inverter PV plant. The acquisition module measures the temperature and solar irradiance on the PV plant in real time and extracts the DC current and voltage observed on the respective array from each MPPT. Sensors mounted on the PV field can be used to measure temperature and solar irradiation. They can also be obtained via a remote database linked to a weather station. Then, the acquisition module feeds these measured data to the detection module. The detection module estimates the DC power that each array should output if no fault occurs using these data from the acquisition module. This module compares the estimated and measured powers and generates an alert signal in correspondence with the arrays that provided a lower power than the estimated one; if the difference exceeds a threshold. Finally, the alert signals are sent to the diagnosis module, which can automatically provide information on the type of fault that occurred.

**Figure 7.** Proposed intelligent system's schematics [103].

The authors of [105] noted that numerous literatures have proposed methods of shadow detection and the reconfiguration of an array. In addition, most of these methods use the voltage, current, and power information to achieve this. The authors saw that monitoring these factors was time consuming and tiresome. Therefore, they presented a novel and effective method of shadow detection for the reconfiguration process, which will help increase the energy production of PV arrays, based on the fuzzy logic and computer vision. The method detects the edge of the object region on the panel from images taken with a camera. Using the background subtraction and object detection method, it then converts the background and foreground image into a grey image format. Object edge detection is performed after the determination of the updated mechanism, which is based on background subtraction and illumination variability. First, dilation and erosion operations are applied to the binary mask in order to determine the object region edge. The image's noise is reduced by these techniques. Following that step, using the Canny edge detection approach, the edges of the objects on the binary image are detected. The final stage in the determination of the subject borders on an image is to use the search and draw contour procedures to determine the related object regions. A relevant pixel region is created for each region whose side information is identified. Finally, the proposed method uses a fuzzy decision-making mechanism to classify the object region as a shadow region utilizing brightness and color distortion values of the object as input parameters. When employed as an input parameter for the reconfiguration operations, the proposed method has a success rate of 98% and increases energy usage performance by roughly 10–15%. The arc defect detection technique described in [106] has a minimal computational requirement, and can function with most of the conventional analog-digital converters (ADC) found in microcontrollers, making it useful in PV applications. When using an inverter rather than a solo device, the algorithm produces better results. This algorithm's integration is also a cost-effective way to detect arc faults and improve the PV system long-term safety. The short- and long-term measurement results are promising, but further long-term experiments are needed to fine-tune the device for phenomena that have yet to be identified. In the proposed method, the detection algorithm uses three indicators, namely frequency analysis, peak detection, and observation of the operating point, as parameters for fault detection. When an arc fault occurs, the indicators display a particular behavior. However, it is not always the same. The signal energy increases slowly at times and quickly at others. When an arc fault occurs in a small plant, the trip is apparent, but not in a large plant. Due to the haziness of the situation, the authors employed fuzzy logic to detect an arc defect. Furthermore, fuzzy logic makes it easier to incorporate the experiences of experts who are unfamiliar with the algorithm. Four sub-detectors, which are followed by the master fuzzy arc fault detector (MFAFD), are created to keep an overview of rules and input variables. These sub-detectors are the peak evaluator sub-detector (PESD), which analyzes all of the peaks and delivers a mass output proportional to the probability, indicating that this peak is from an arc fault; window near sub-detector (WNSD), which analyzes the change signal energy over a short timeframe; window wide sub-detector (WWSD), which analyzes the long-term signal energy and can be used when there is no abrupt signal energy growth; and power analyzer sub-detector (PASD), which supervises the power change. Outputs of the four sub-detectors serve as input for the MFAFD. The MFAFD outputs a number between 0 and 1, which represents the mass for the arc fault probability. An arc fault is detected if this probability exceeds a predefined threshold. The authors of [107] presented a method for detecting increases in series resistance using a fuzzy classifier that can distinguish between the increasing series losses and partial shadow situations for resistances greater than 400 W/m². As shown in Figure 8, an optional shadow detection algorithm acting before the increased series losses detection system, which could improve the detection accuracy of the system, is also implemented in the diagnostic system. This strategy is especially significant, since the increased series losses and partial shadow circumstances are difficult to discern, as they diminish a PV system's peak output and fill factor. Rather than the controlled laboratory circumstances, the study

focuses on estimating the increased series resistance in the field. The method has been tested using experimental measurements. In addition, it has shown good detection rates across a wide range of irradiance levels, as well as in the presence of diverse sizes and patterns of partial shadows. Moreover, the authors showed that a dedicated partial shadow detection algorithm, implemented in the diagnostic system and functioning prior to the higher series losses detection method, improves the overall system's detection accuracy.

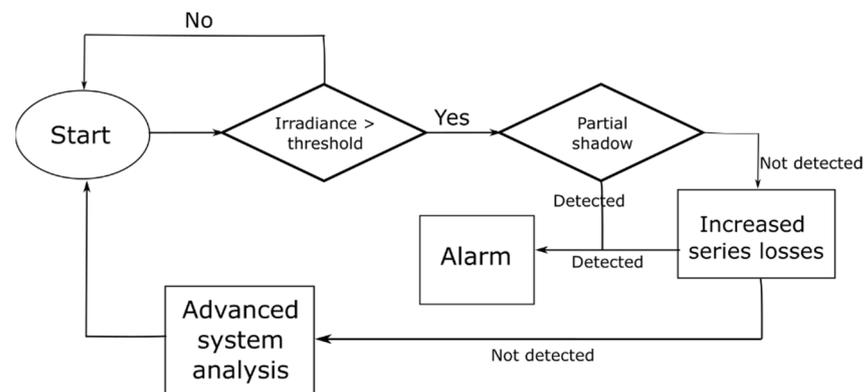


Figure 8. Proposed diagnostic system of partial shadow detection, increased series losses detection, and advanced system analysis and monitoring for the PV module [107].

Under low irradiance conditions, the DC side short-circuit faults in PV arrays consisting of multiple PV panels connected in a series/parallel configuration are nearly undetectable, especially when the MPPT algorithm is in use. In addition, if they go undetected, these faults can significantly reduce the output energy of PV systems, damage the panels, and potentially cause fire hazards. To avoid this, the authors of [108] present a fault detection scheme based on a pattern recognition approach that uses a multi-resolution signal decomposition technique to extract the necessary features, which is then used by a fuzzy inference system to assess whether a defect has occurred. PV array output volts and currents, as well as solar irradiation, are the system's inputs. The multi-resolution signal decomposition technique is used to extract four unique features depending on the above inputs. Following that step, the retrieved characteristics are supplied into the fuzzy inference system, which generates a scalar number based on carefully built membership functions and the associated rule base. Decisions are based on these results. The amount of this output determines whether a line-to-line, line-to-ground or none-of-the-above fault has occurred. In simulation and experiment-based case studies, the performance of the proposed method is demonstrated. Defect identification becomes increasingly difficult when the percentage of fault impedance or mismatch declines. The program also revealed one unintentional operation for a typical case, in which the irradiance had changed drastically within a short period of time, which is unusual in real-world systems. The promising performance of the proposed algorithm is supported by the experimental results. Line-to-line and line-to-ground faults are usually detected using this method. Open-circuits and hot-spot heating, for example, are two further types of PV array faults. As a result, the applicable algorithms for a more exact classification of additional PV concerns could be applied in parallel with or later phases of the detection system. Fuzzy logic-based algorithms are presented in [109,110] to detect the malfunctioning PV modules and partial shadowing circumstances that influence the DC-side of grid-connected PV systems. The authors' algorithm is made up of six layers that work in a sequential order. Input parameters make up the first layer (solar irradiance and module temperature). The theoretical performance analysis of the grid-connected PV system is generated using the LabVIEW virtual instrumentation program in the second layer. The power and voltage ratios are determined in the third layer, and high and low detection limits are set in the fourth layer, which is utilized to apply the 3rd-order polynomial regression model to the power and voltage ratios. The fifth layer contains the input parameters of the examined grid-connected

PV systems, as well as the 3rd-order polynomial detection restrictions. If the measured voltage ratio vs. the measured power ratio is outside of the detection limits, the data will be processed by the sixth layer (which is the last layer), which contains the fuzzy logic categorization system. The suggested method's innovation is proved by the fact that it provides a simple, dependable, and fast fuzzy logic classification system that can be employed with a variety of grid-connected PV systems. The algorithm is also unique in that it is based on fluctuations in the voltage and power of the grid-connected PV system. Few fault diagnosis techniques are capable of being implemented on integrated circuits, according to Sufyan Samara and Emad Natsheh's study [65], and these procedures require expensive and complicated hardware. The authors introduced a unique effective and implementable defect diagnosis approach based on the AI nonlinear autoregressive exogenous NARX neural network and Sugeno fuzzy inference to try to solve this problem. The program employs the Sugeno fuzzy network to isolate and classify errors in PV systems. The NARX network is used to estimate the PV system's maximum output power based on the real-time measured output and surrounding conditions, which is then used by the fuzzy inference algorithm to detect and categorize errors that may develop in the PV system. The algorithm has been demonstrated to work on a low-cost microcontroller. The suggested algorithm will be able to detect a variety of flaws in the PV system, including open- and short-circuit degradation, faulty MPPT, and PS issues. Furthermore, the proposed algorithm can capture non-linear patterns between predictors, such as radiation and temperature, as well as other non-linear correlations of patterns between predictors, to calculate the precise moment of maximum power for the PV system. The actual sensed PV system output power, anticipated PV system output power, and sensed surrounding conditions are all required for fuzzy inference. Using an AI NARX-based neural network, the PV system's output power is projected. The authors concluded that the ability of the proposed method to efficiently diagnose several PV system faults is an important step in achieving a complete system that can diagnose large PV system faults.

Table 5. Summary of neuro-fuzzy-based methods.

Authors	Reference	Year	Contribution
P. Ducange, M. Fazzolari, B. Lazzerini, F. Marcelloni	[103]	2011	Intelligent system for the detection of PV system faults based on TSK-FRBS
K. Karakose, K. Firildak	[105]	2014	Shadow detection system based on fuzzy logic using the obtained PV array images
B. Grichting, J. Goette, and M. Jacomet	[106]	2015	Arc fault detection algorithm in PV systems based on the cascaded fuzzy logic
S. Spataru, D. Sera, T. Kerekes, R. Teodorescu	[107]	2012	Increased series losses detection in the PV array based on fuzzy inference systems
Z. Yi, A.H. Etemadi	[108]	2017	Detection of PV system faults based on the multi-resolution signal decomposition and fuzzy inference system
M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales	[109]	2017	PV system diagnostic technique based on the six layer detection algorithm
M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales, P. Mather	[110]	2017	Fault detection in PV systems based on the theoretical curves modeling and fuzzy classification system
S. Samara, E. Natsheh	[65]	2020	Using the NARX network and linguistic fuzzy rule-based system to diagnose PV panel faults

4.6. Wavelet-Based Methods

In this subsection, the wavelet-based methods are reviewed. A summary of the method contributions presented in the literature is shown in Table 6.

The authors in [111] proposed a system that was designed and tested at various load currents, DC source voltages, and arc lengths to analyze the impact of each parameter on the DC arc, with the goal of providing a thorough understanding of arc behavior that could occur in DC networks. Based on the findings, a detection technique that uses both the time domain and time frequency domain characteristics to distinguish between the DC arc fault and normal conditions has been presented. A simple calculation approach for the arc current variation is utilized to portray the chaotic and dynamic nature of the DC arc physical process. The current variation can indicate different stages of arcing and is used to recognize DC arc faults. Moreover, wavelet decomposition is applied to the arc current signal. The normalized RMS value of wavelet decomposition coefficients demonstrates the ability to detect the arc fault. Finally, a comprehensive DC arc detection algorithm is created using the current change from the time domain analysis and the normalized RMS value from the wavelet analysis. A Texas Instrument's Digital Signal Processors (TI DSP) chip, which has a clock frequency of 150 MHz and has gained popularity in power electronic applications, is used to implement the detection method in the hardware. In the proposed method, the maximum and minimum currents are updated every time new data is inputted at each timer cycle, and the difference in the current is registered at the n th cycle. The value of n is determined by the length of the time window and sampling frequency. Detection accuracy is enhanced if a longer time window is achieved, as it makes the RMS versus time waveform smoother. A long time window, however, leads to a longer time response. When the amount of data points is excessive, the actual implementation must be adjusted to ensure that the calculation load is within the capabilities of the available microprocessor. The wavelet coefficients are calculated by sending a signal through low-pass and high-pass filters, as well as down sampling it by two. The filters are dependent on the used wavelet and are consistent throughout the entire level decomposition. As a result, the detection algorithm can be implemented on a microprocessor in the time domain using convolution and down sampling. The filtering procedure for the first-level DWT to obtain *Coefl1* is carried out every two timer intervals, for which two new data points are available at this time. The second-level DWT coefficients (*Coefl2* and *Coefh2*) are calculated whenever two new coefficients from the first-level DWT are received. As the filtering operation was performed for every timer interval instead, this method cuts the calculation load by half. In the main function, the ultimate decision is taken. The main function's flag changes to 1 at the conclusion of each time window. Only two square roots, one division, and two comparisons are performed in the main function, which takes 80 clock cycles in total, according to the benchmarks of the C28x floating point unit fast RTS library. With a clock frequency of 150 MHz, the main function's computation time at the conclusion of each $T_{sw} = 25$ ms is less than 1 μ s, which is negligible. As a result, the detection time is mostly determined by the length of the time window. The results of the studies reveal that the detection algorithm is capable of delivering an alarm in a timely manner following the initiation of an arc fault. Typical operations, such as load changes, can also generate nuisance tripping, which can be reduced with this detection system. In [112], a modified wavelet-based technique termed the wavelet packet transform (WPT) was developed for the detection of diverse disturbances caused by faults in grid-connected solar PV systems. The study continued to evaluate the proposed WPT approach to methods based on the ordinary wavelet transform (WT) under various operating settings. In addition, qualitative and quantitative evaluations suggest that the WPT outperforms WT in terms of detection performance. The WPT uses a set of low-pass and high-pass filters to breakdown the signal retrieved at PCC. It gives both approximate and precise coefficients. Moreover, it extensively decomposes both components in order to determine the signal's frequency agreements. The breakdown procedure carried out in both components is what gives them their value. Wavelet transforms, on the other hand, do only a one-sided decomposition, segmenting only the low-pass frequency components and not the high-pass frequency components. When WT is subjected to noisy or transitory environments, this feature can compromise its performance. In [113], wavelet transformations were used to offer an online

fault detection method for power conditioning systems (PCS). Switch open faults and over harmonics are detected using a multi-level decomposition wavelet transform approach. Using the normalized standard deviation of the wavelet coefficient, a quick and accurate diagnostic function is also achievable, allowing the suggested method to detect islanding conditions. Simple calculations (a time correlation generated by sequential multiplication and addition) and exact diagnostic capabilities of fault identification with good simulation outputs to check and evaluate its claims characterize the method's algorithm. The multi-level decomposition wavelet transform provides the method's straightforward calculation characteristic. At each wavelet tree level, the algorithm extracts wavelet coefficients from the measured signal, and errors are discovered and categorized using the wavelet coefficient changes. Using a three-level MLD tree, the fault detection technique was created for PCS fault scenarios. Switch open and over harmonic are the two scenarios considered. The PCS uses a semiconductor switch, such as a field effect transistor (FET) or an IGBT, to convert DC solar voltage to AC grid current. The switch open can be attributed to a switching device failure, whereas the over harmonic can be attributed to a controller or sensor failure. The cases of switch short faults are not taken into account here. The system is protected by the over-current limiting function or melting fuse when the switches are in short fault, and the PCS ceases running. The PCS current has a distorted waveform when the switches are open faulted, and it continues to provide high order harmonics to the grid. An UP and DOWN switch failure in the inverter bridge might be categorized as an open switch problem. The authors of [114] used the discrete wavelet transform (DWT) to analyze the traced I-V curve of a residential PV system and define these coordinated points in the related diagnosis effort. The DWT was utilized to implement the fault diagnosis of residential PV systems as a preprocessing tool. It enables feature extraction through signal decomposition and noise reduction. The reduced short-circuiting current of partially shaded cells is represented by the vertical height or current in a PV string identified by the DWT method, whilst the horizontal or voltage distance from the VOC to the inflex is connected to the number of bypassed modules. The approach is divided into two sections, namely passive and active. In the passive diagnosis section, a residue signal is generated by comparing the measured PV power signal and simulated model in real time to monitor the alarm signal and abnormal condition in the system, using the model base fault diagnosis technique. After the manifest and certainty error signals have been determined, the flash test is used as the active and second portion of the test procedure. During this phase, the step load is separated from the PV and power generation, and the MPPT mode of the inverter is interrupted. The I-V curve of the PV array is tracked and logged by modifying the inverter switching pattern for a deeper inspection and interpretation. The model provided in [115] for the defect diagnosis of PV arrays was improved using improved wavelet neural networks, wavelet neural networks, and back propagation neural networks. The training technique now includes a Gaussian function, which is utilized as an activation function, an additional momentum mechanism, and an adaptive learning rate method. The conclusion is taken from simulation findings that the proposed technique in this study is capable of efficiently diagnosing the PV array problem with good performance accuracy, convergence time, and stability under the identical conditions of the network input and desired output. The proposed fault diagnosis algorithm is summarized in Figure 9. Four PV system fault types are diagnosed using the model, namely short-circuit, open-circuit, abnormal degradation, and partial shading. There are five network output layer variables since the system requires the precise diagnosis of four types of problems and no fault condition. The selection of the number of hidden layer nodes in a neural network is a difficult topic with no theoretical foundation to follow. The number of hidden layer nodes is critical to the network's success. If the number is too little, we may not acquire a network from training, implying that the network's robustness is poor and its anti-noise ability is weak, making it unable to recognize models that have never been seen before. If the hidden layer's node number is too big, the learning time will be too long, and the error will not be minimal. Furthermore, there could be an issue with overfitting. As a result, based on

an empirical equation, the trial and error method is commonly employed to identify the appropriate number of concealed nodes.

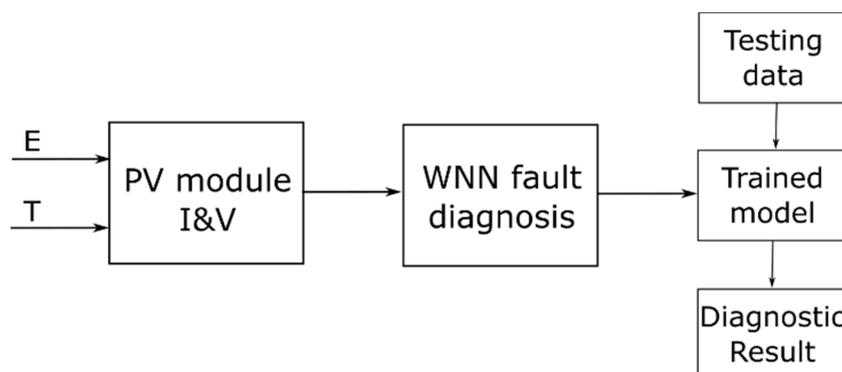


Figure 9. Proposed fault diagnosis method based on the wavelet neural network [115].

In [116], a wavelet optimized exponentially weighted moving average (WOEWMA) monitoring technique based on the principal component analysis (PCA) was created. The presented monitoring approach combines the advantages of exponentially weighted moving averages, MOO, and wavelet representation. MOO is used to tackle the challenge of determining the best strategy for minimizing both the missed detection rate (MDR) and the false alarm rate (FAR). Apart from the decorrelation of auto-correlated observations, the wavelet representation increases the monitoring performance by lowering MDR and FAR, as well as obtaining the exact deterministic features. The proposed method is appropriate for real-time implementation due to its quick calculation time. This approach is robust in a range of faulty conditions, including single and multiple (cascade and simultaneous) occurrences, and can detect faults in both the DC and AC conversion sections. The proposed method has several advantages, including dynamic multiscale representation to extract accurate features and de-correlate auto-correlated measurements; the PCA technique to model PV systems; the EWMA of two parameters in OEWMA are optimized using the multi-objective optimization; and the WOEWMA chart can detect smaller fault shifts, thereby improving the PV model monitoring.

Table 6. Summary of wavelet-based methods.

Authors	Reference	Year	Contribution
X. Yao, L. Herrera, S. Ji, K. Zou, J. Wang	[111]	2014	Detection of DC arc fault based on the characteristic study and time-domain discrete wavelet transform
P.K. Ray, A. Mohanty, B.K. Panigrahi, P.K. Rout	[112]	2018	Fault analysis in solar PV system based on the modified wavelet transform
I.S. Kim	[113]	2016	Online fault detection algorithm in PV system based on the wavelet transform
M. Davarifar, A. Rabhi, A. Hajjaji, E. Kamal, Z. Daneshifar	[114]	2014	Partial shading fault diagnosis based on the discrete wavelet transform
X. Li, P. Yang, J. Ni, J. Zhao	[115]	2014	Improved wavelet neural network algorithm-based method of fault diagnosis in PV array
M. Mansouri, A. Al-khazraji, M. Hajji, M.F. Harkat, H. Nounou, M. Nounou	[116]	2017	Wavelet optimized EWMA-based fault detection method for the PV system application

4.7. Other Methods

This subsection groups all the other reviewed faults that were not classified in any of the six groups. A summary of the contribution of the reviewed methods is presented in Table 7.

The approaches, that use graph-based semi-supervised learning (GBSSL) for fault identification and diagnosis in PV arrays, are available here [13,117]. The authors of [117] presented a sound technique based on GBSSL for recognizing, classifying, locating, and fixing errors in PV arrays, which was improved by expanding the diagnostic space of the GBSSL algorithm and adding more class labels. The model detects and locates a failure, temporarily isolating the system in order for it to continue to operate normally until the problem is resolved. The authors adjusted the way the data were normalized in order to increase the system's ability of finding the unlearned defects, and they were able to detect faults that the algorithm could not detect at first. The functionality of the system to feed the required energy after the defect was eliminated was tested using an interleaved boost converter, and it was discovered that the maximum voltage loss in a standard condition is 1 V, demonstrating the model's high efficiency. The proposed model in [117] was an improvement of a previous model proposed in [13] by the addition of the fault location element and higher fault detection and classification accuracy. The early stage of the algorithm starts with the extraction of a limited number of labeled data with their class labels serving as initial value data and fed into the model. Both the faulted and fault-free data are included in this dataset. To achieve the fault-prone labeled data, faults are purposefully introduced into the system, and the parameters are measured and stored for each data. The data are measured and stored in the fault-free mode in the same way as it is in the normal mode. An important advantage of the proposed GBSSL method is that it only requires a small amount of data for the learning process. To detect PV system faults, the current of all rows of solar cells, as well as the overall voltage of the panel, are constantly measured at any time (this is accomplished by installing 10 current sensors in each row of cells and one voltage sensor at the PV system's output points). As a result, each dataset contains 11 parameters, 10 of which are currents flowing across the cell's rows and one of which is the system's total voltage. In the modeling process, the parameters required to implement the GBSSL algorithm are entered. On the other hand, the authors of [118] proposed an intelligent defect diagnostic approach for PV arrays based on a kernel extreme learning machine (ELM) optimized simulating annealing (SA) algorithm with an improved radial basis function (RBF). Short-circuit faults, aging faults, and shadow faults are among the faults discovered by the proposed method. The results obtained were above 90%, showing that the SA-RBF-ELM fault classification is accurate and stable. According to the acquired simulation findings, the suggested model has three key advantages: (1) The optimal fitness value of the PV array and the model parameters used, as one of the characteristic factors of neural network learning, considerably increase the fault detection accuracy of the four fault kinds described; (2) the RBF-ELM kernel function has strong learning and classification capabilities, making it suitable for detecting and classifying PV array faults; and (3) the SA algorithm can quickly optimize the parameters of the RBF-ELM fault diagnostic model, significantly improving the RBF-ELM model's training accuracy and testing precision. The algorithm proposed is a derivative of the basic ELM algorithm, with kernel function limit learning machine features, which improve its ability to solve the problem of regression prediction leading to higher accuracy and faster calculations. The RBF-ELM model contains a regularization coefficient (C) and radial width (α), which affects the algorithms' performance. The RBF-ELM fault model's training accuracy is employed as the optimization objective function, and the coefficients C and α are the parameters that must be tuned. The simulated annealing approach is then utilized to optimize the parameters of the RBF-ELM fault model, resulting in optimal training and test accuracy for each time. A diagnostic technique for PV systems was developed in [119] using the learning method to take each PV site's condition into account. The technique employs the diagnostic criteria database to analyze the data acquired from the PV system.

The special features of the proposed technique include updating the diagnostic criteria, making it possible to detect normal or abnormal operating conditions of a PV system; the detection of shadow on modules and the pyranometer using the sophisticated verification (SV) method [120]; and the maintenance advice provided by an expert system according to the precise diagnosis. The ratio of acquired data to reference data is calculated to diagnose the system's normality or abnormality. The ratio approaches "1" when actual and average meteorological data are close. For example, when the summer generated power ratio is "1" and the winter generated power ratio is "0.7," a winter shadow or snow on the modules is assumed. The criterion for diagnosis in this situation is "1". The contribution of the proposed method is highlighted in its features, as follows: By updating the diagnostic criteria, it is now feasible to diagnose the normality or abnormality of PV systems, while taking into account the PV system's characteristics, as well as the climate; where a shadow appears on the modules or pyranometer is determined using the SV approach and hourly data analysis; and maintenance recommendations are also given based on the diagnosis outcome. The simulation results of the proposed technique suggest that it offers quick and proper maintenance advice within a short detection period. In [121], a simple short-circuit and open-circuit fault detection approach for PV systems was suggested based on the evaluation of three coefficients. The suggested technique has two steps. First, an offline simulated model for extracting the variation boundaries of the three coefficients for each faulty operation. Second, an online comparison model for comparing real measured coefficients to the simulated coefficients from the offline step. Three coefficients have been established for each fault type in order to detect and diagnose both short-circuit and open-circuit faults, namely the current coefficient, the voltage coefficient, and the power coefficient. The offline step is aimed at extracting the three coefficients' variation boundaries for each type of fault. In order to achieve this, three other operations are conducted. By bringing the detected parameters to a PSIM/MATLAB co-simulation, you may simulate both the healthy and flawed scenarios under a few climatic situations. For each simulated instance, the goal of this stage is to extract a few MPP coordinates. Based on the given equations, determine the three coefficients for each fault situation, then by adding a $\pm 2\%$ offset to the three derived coefficients, you may extract the variation boundaries for each defective type. For the online step, using the various sensors, both the meteorological conditions and MPP values may be detected and monitored during the actual operation of a PV system. The three actual onsite coefficients will be calculated using these measures. Finally, a comparison is made between the real onsite coefficients and the variation boundaries of each faulty case that was previously stored during the offline process. To conclude, the faults detection task will be carried out based on the real onsite monitored power coefficients measured, in a way that if their value exceeds a set threshold, a DC side fault alarm will be triggered. In addition, the faults' type will be determined by comparing the three real onsite coefficients with the variation boundaries of each simulated faulty case. The proposed method is straightforward, efficient, and does not necessitate a large amount of training data. The authors of [71] presented a data-driven anomaly detection and classification system that can accurately detect and categorize a wide range of PV system anomalies. The method consists of two stages. First, the local context-aware detection (LCAD), which is a hierarchical context-aware anomaly detection using supervised learning, and is aimed at identifying possible anomalies in PV strings with current characteristics that are different from the other PV strings under similar environmental conditions. Second, the remote context-aware detection (RCAD), which is a hierarchical context-aware anomaly detection using supervised learning, and is aimed at identifying possible anomalies in PV strings with current characteristics that are different from solar PV farms and benefit from a combination of LCAD and GCAD to detect anomalies at the string level. First, the domain-specific features are designed. To reduce computation complexity and increase classification performance, the multimodal properties are carefully generated and extracted. Then, with the purpose of developing an accurate classification model that is suitable for specific categorization situations, a

multimodal model training technique is constructed. The effectiveness, robustness, cost-, and computing efficiency of the suggested strategy are proved by the results of trials conducted over time. The proposed method has the following advantages: A more robust method against irradiance and weather variations that can accurately detect different anomalies without pre-labelled data; 90.2% detection accuracy for the top 100 anomalies that are otherwise nearly undetectable under low irradiance or weather with high cloud cover; the use of SCADA data to classify commonly occurring anomalies at the plant level; and cost- and computation-efficient as it uses readily available data of existing PV systems. Numerous machine learning-based fault detection methods have the following problems, according to the authors in [122]. Fault diagnosis performance is limited due to the insufficient monitored information. Moreover, fault diagnosis models are inefficient to train and update, and labeled fault data samples are difficult to obtain by field experiments. The authors proposed a method with the aim of overcoming these problems and three features were addressed. The first is based on important points and model parameters collected from I-V characteristic curves and environmental factors that are observed. An effective and efficient feature vector of seven dimensions is proposed as input of the model. The second is an emerging kernel based on extreme learning machine (KELM), which features extremely fast learning speed and good generalization performance, utilized to automatically establish the fault diagnosis model. The Nelder-Mead simplex (NMS) optimization method is employed to optimize the KELM parameters, which affect the classification performance. The final aspect is an improved accurate SIMULINK-based PV modeling approach for a laboratory PV array to facilitate the fault simulation and data sample acquisition. There are six steps leading to the establishment of the proposed model, as shown in Figure 10.

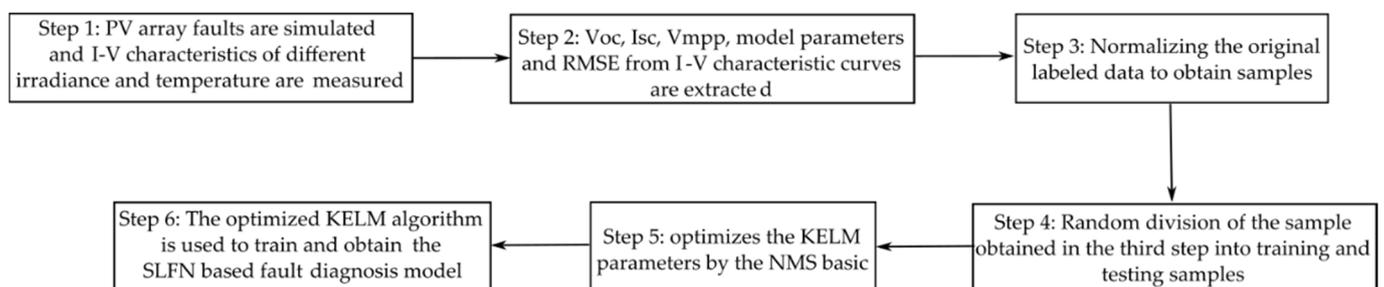


Figure 10. Flowchart on the establishment of the proposed fault diagnosis model [122].

The data samples for each fault condition should cover a wide range of operational irradiance and temperature, in order to make the fault diagnosis model suitable for a variety of operating settings. To begin, certain SIMULINK simulation experiments were used to obtain labeled data samples of normal and problematic situations. Then, on the real laboratory PV array, some field experiments were conducted to achieve some experimentally labeled data samples. Finally, the fault diagnostic model is established using the optimized KELM, which is evaluated and analyzed using both simulated and experimental data samples with known fault kinds. The proposed KELM-based fault detection model is promising in real-time applications due to its exceptionally fast learning speed, simplicity, and high generalization performance. The authors are attempting to apply the fault diagnosis model in digital signal processor (DSP) based embedded real-time systems, in conjunction with an integrated rapid I-V tester that is currently in development. The authors of [123] presented outlier detection rules based on instantaneous PV string current monitoring for failure detection. It is a command to monitor PV functioning and discover faults that may go undetected by overcurrent protection devices (OCPD). Three outlier identification rules were devised and compared by the authors, namely the three-sigma rule, Hampel identifier, and boxplot rule. Weather measurement or model training are not required with the suggested strategy. The Hampel identifier performs well in

cases with extremely high contamination levels (33.3% in this investigation), while the boxplot rule performs better under PV faults in cases with relatively high contamination levels (12.5% in the case of this study). The model's reliability improves as the number of PV measurements rises. Despite the fact that the outlier identification methods in this study are based on PV-string level measurements, the authors claim that the proposed approaches should be straightforward to implement with minor modifications on any PV installation level. If the assumption is made that the solar irradiation is identical on the same PV level, the outlier rules can be applied on the PV-module or sub-array level, for example. Aside from the PV string current, the measurement could include PV insulation impedance, output power or energy yield, all of which are commonly used PV metrics. This may provide extra flexibility to fault detection methods. To overcome the limitations of conventional wired monitoring systems, such as physical constraints during data cable laying, high installation and maintenance costs, and reduction in the system lifespan due to the over exposure to extreme weather conditions, a Zigbee-based wireless monitoring system was developed in [124] to replace the conventional systems for online monitoring of parameters, such as temperature, irradiation, PV power output, and grid inverter power output, in grid-connected PV system applications. Moreover, it is equipped with a control function for remote system monitoring and a user-friendly web application, in order for the monitored data to be easily accessible via the Internet. Although the simulation results were satisfactory, the authors pointed out the limitations of the proposed method. These limitations include: (1) The proposed method is location specific. Therefore, before implementing the system in other locations with significantly different weather conditions, the weather factor needs to be taken into consideration. (2) The program used was developed based on the available software and programming language familiar to the authors. In [125], a failure diagnostic algorithm based on an online distributed monitoring system of a PV array of Zigbee wireless sensors network and a genetic algorithm optimization based BP neural network was investigated. The Zigbee wireless network system monitors each module's output current, voltage, and irradiation, as well as the environment's temperature and irradiation. In addition, a simulation PV module is set up, based on which typical problems are simulated and fault training samples are obtained. The fault sample data are then utilized to build and train a generic algorithm optimized BP neural network diagnosis model, which is subsequently used to detect four different PV array operating states (normal, abnormal aging, short-circuit, shadow). Since an open-circuit problem is noticed during the data collecting phase, it is not considered one of the diagnosis model's outputs. According to the simulation data, the proposed deflection system has a high level of accuracy. Operators or managers can log in to verify the parameters of each PV module and use the designed mechanism to quickly discover the problematic PV module. In relation to the PV energy conversion system (PVECS), the authors of [126] presented a fault detection system using a fractional-order color relation classifier. The output power degradation is used to monitor the physical circumstances associated with changes in the circuitry of a PV array, such as grounded faults, mismatch faults, bridged faults between two PV panels, and open-circuit faults, using an electrical inspection method. The over-current and ground fault prevention devices can also be used to isolate failures on the AC side. As a result, the grid connection side fault impact can be reduced. Iterative calculations are not required to update the parameters of the inference model in the flexible and inferential model. As a result, it can handle the complexity of an adjustable mechanism in a relatively short design cycle. Embedded system approaches can then be used to implement the proposed detection model. The suggested approach can detect normal conditions, mismatch faults, and four common electrical defects on the DC side, according to the simulation results. For solar radiation of 0.4–1.0 kW/m² and temperatures of 25–40 °C, the suggested detection model has an average accuracy of 88.23% in identifying the fault under low/high solar radiation and various temperatures. The authors of [127] presented another interesting fault detection and diagnosis method based on a laterally primed adaptive resonance theory (LAPART) neural network. It is a

low-cost way of automatically detecting and diagnosing PV system issues. The LAPART algorithm was taught how to detect fault states using real-world data that were classified as normal system behavior. The algorithm was then given new data and three-fault data points for an initial test. The system was given synthetic data to examine its performance over a statistically significant month-long dataset, and it was able to correctly identify flaws within the dataset. The LAPART algorithm's accuracy is determined by its ability to deliver a high likelihood of detection, while reducing false alarms. The number of true positive values generated by the FDD process is compared to the total number of actual positive values to determine the likelihood of detection. The LAPART architecture combines two fuzzy adaptive resonance theory (ART) algorithms to build a system for predicting outcomes based on the learnt associations. The single fuzzy ART algorithm's fundamental equations include category selection, match criterion, and learning. The goal is to create the optimal template matrix for the provided dataset. The approach employs category selection to discover the existing template matrix that best matches the provided input. In addition, for fast learning applications, the free parameter is frequently set to 10⁻⁷. The match criterion then checks to verify if the template matrix and input that is compared fulfill the user-defined vigilance parameter criterion. Depending on the level of intricacy requested, the vigilance free parameter can range from 0 to 1. A high vigilance value of 0.9, for example, yields high complexity but limited generality, whereas a low parameter of 0.5 yields the opposite. Finally, if it passes, the template is changed to reflect what has been learned. The LAPART algorithm is created by linking the two fuzzy ARTs (FARTs), which is seen graphically in Figure 11. The L matrix, which connects the A and B templates, connects the A and B FARTs. Each FART has its own set of vigilance settings, and inputs are delivered to both the A and B sides at the same time during the learning process. The A and B sides work together to generate and update the templates, while also forming links. Testing inputs are only applied to the A side after the training is complete, allowing them to resonate with the already acquired templates. The L matrix's relationships are then used to link with the B side and generate the prediction results.

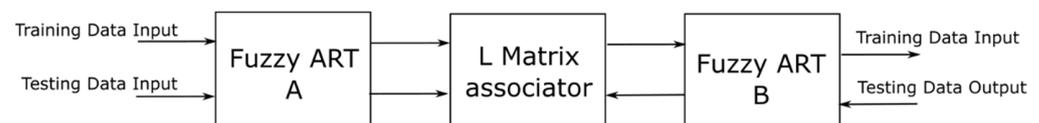


Figure 11. Creation of LAPART algorithm through linking two fuzzy ARTs [127].

The approach provided in [128] is based on monitoring the PV array's output power and is suitable for low irradiance, high impedance, and low mismatch fault circumstances. The irregularity of the time series of the normalized fault-imposed component of PV power is measured using entropy-based complexity as the fault detection criterion. Weather disturbances and partial shade can cause many array faults, which the proposed method can identify. It is applicable to both grid-connected and islanded PV systems and does not require a training set or prior knowledge of the PV array. Moreover, it is an economical strategy as it does not require costly sensors, relying only on the central IED to process the PV voltage and current measurements. The irregularity of the fault imposed power time series is measured by the sample entropy. The study uses the sample entropy-based complexity as the PV array fault detection index (FDI), since the complexity of time series data more effectively captures the behavior of a nonlinear system. The fault-imposed component of the PV array power is zero during the normal operation. As a result, the time series of moving data windows is regular. Therefore, SampEn is equal to zero. The fault-imposed power samples in each moving data window of N points are not identical when the solar irradiance or temperature varies, but they are fairly close together, since weather disturbances are not severe. As a result, FDI will be near-zero in this situation. When a fault develops in the PV array, however, the fault-imposed power samples rapidly shift. As a result, the normal fault-imposed power samples differ dramatically from the post-fault ones. Therefore, non-repetitive patterns can be found in the moving data windows during

the initial milliseconds of fault transients, and the estimated SampEn is not zero. The issue is recognized when FDI rises to a non-zero value. It may be concluded that FDI for normal occurrences is approximately zero, whereas FDI for fault events is non-zero. A defined threshold is used to discriminate between non-zero FDI values under fault situations and near-zero FDI values during weather disturbances and partial shadings. As a result, determining the fault detection threshold is simple. It can considerably reduce nuisance tripping. Several experiments that were carried out validate the proposed fault detection algorithm's simplicity, sensitivity, scalability, resilience, and adaptability.

Table 7. Summary of other methods.

Authors	Reference	Year	Contribution
Y. Zhao, R. Ball, J. Mosesian, J. De Palma, B. Lehman	[13]	2015	Graph-based semi-supervised learning algorithm for PV fault detection and classification
H. Momeni, N. Sadoogi, M. Farrokhifar, H.F. Gharibeh	[117]	2020	Graph-based semi-supervised learning method of PV array fault diagnosis and correction
Y. Wu, Z. Chen, L. Wu, P. Lin, S. Cheng, P. Lu	[118]	2017	SA-RBF kernel extreme learning machine-based intelligent fault diagnosis for PV array
Y. Yagi et al.	[119]	2003	Diagnostic technology and expert system for PV system using the learning method
E. Garoudja, K. Kara, A. Chouder, S. Silvestre, S. Kichou	[121]	2016	Short- and open-circuit fault detection in PV system based on the offline and online evaluation of measured coefficient
Y. Zhao, Q. Liu, D. Li, D. Kang, Q. Lu, L. Shang	[71]	2019	Hierarchical context-aware anomaly detection using the unsupervised learning and multimodal anomaly classification method
Z. Chen, L. Wu, S. Cheng, P. Lin, Y. Wu, W. Lin	[122]	2017	Optimized kernel extreme learning machine and I-V characteristics based intelligent fault diagnosis in PV array
Y. Zhao, B. Lehman, R. Ball, J. Mosesian, J.F. De Palma	[123]	2013	Outlier detection based rules for solar PV array fault detection
F. Shariff, N.A. Rahim, W.P. Hew	[124]	2015	Online monitoring of grid-connected PV system based on the Zigbee data acquisition system
H. Lin, Z. Chen, L. Wu, P. Lin, S. Cheng	[125]	2015	Online monitoring and fault diagnosis of PV array based on the genetic algorithm optimized BP neural network
C.L. Kuo, J.L. Chen, S.J. Chen, C.C. Kao, H.T. Yau, C.H. Lin	[126]	2017	PV energy conversion system fault detection using the fractional order color relation classifier
C.B. Jones, J.S. Stein, S. Gonzalez, B.H. King	[127]	2015	Photovoltaic system fault detection and diagnostics using the laterally primed adaptive resonance theory neural network
A. Khoshnami, I. Sadeghkhani	[128]	2018	PV array fault detection based on sample entropy

5. Discussion

This study presented a comprehensive review of PV system fault detection and diagnosis techniques that are based on artificial intelligence and machine learning. Conventional fault detection and diagnosis methods, which equip PV systems with overcurrent protec-

tion devices and ground fault detection interrupters, are not sufficient enough to detect certain faults due to low irradiance conditions, nonlinear output characteristics, maximum power point tracker of PV inverters or high fault impedances. This led to the need for more intelligent fault detection and diagnosis methods to replace the conventional methods, in order to improve the PV systems operational efficiency and safety. AI-based methods, which are still currently explored and improved have been found to be the alternative to conventional methods. This paper's contribution outlines the main features of reviewed AI-based methods and the effectiveness of PV fault detection and diagnosis applications. The reviewed methods mostly adopt ML models, such as neural networks, wavelets, fuzzy logics, decision trees, support vector machines, graph-based semi-supervised learning, regression, etc., in order to develop models and algorithms that are trained to learn the relationships between input and output parameters of PV systems. The effectiveness of these methods depends on their ability to detect a fault and pinpoint its location in the shortest possible time; their relative affordability; and ease of use. Of note, there are currently fewer literatures in this area of PV application compared to the other areas, since the topic has only recently been explored, as evident in the oldest paper we could obtain, which dates back to only about 15 years.

Author Contributions: Conceptualization, A.A. and C.F.M.A.; methodology and formal analysis, A.A., C.F.M.A. and M.G.; writing—original draft preparation, A.A.; writing—review and editing, C.F.M.A. and M.G.; supervision, C.F.M.A. and M.G.; funding acquisition, A.A. and C.F.M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (No. 88887.514132/2020-00).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors kindly thank ENERQ-CTPEA Centro de Estudos em Regulação Qualidade de Energia for their support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Erdiwansyah; Mamat, R.; Sani, M.S.M.; Sudhakar, K. Renewable energy in Southeast Asia: Policies and recommendations. *Sci. Total Environ.* **2019**, *670*, 1095–1102. [[CrossRef](#)]
2. Emodi, N.V.; Chaiechi, T.; Alam Beg, A.R. The impact of climate variability and change on the energy system: A systematic scoping review. *Sci. Total Environ.* **2019**, *676*, 545–563. [[CrossRef](#)]
3. Wilberforce, T.; Baroutaji, A.; Soudan, B.; Al-Alami, A.H.; Olabi, A.G. Outlook of carbon capture technology and challenges. *Sci. Total Environ.* **2019**, *657*, 56–72. [[CrossRef](#)]
4. Wilberforce, T.; El Hassan, Z.; Durrant, A.; Thompson, J.; Soudan, B.; Olabi, A. Overview of ocean power technology. *Energy* **2019**, *175*, 165–181. [[CrossRef](#)]
5. Osmani, K.; Haddad, A.; Lemenand, T.; Castanier, B.; Ramadan, M. A review on maintenance strategies for PV systems. *Sci. Total Environ.* **2020**, *746*, 141753. [[CrossRef](#)] [[PubMed](#)]
6. Varella, F.K.O.M.; Cavaliero, C.K.N.; Silva, E.P. Regulatory incentives to promote the use of photovoltaic systems in Brazil. *Holos* **2012**, *3*, 15–29. [[CrossRef](#)]
7. Block, P.A.B.; Salamanca, H.L.L.; Teixeira, M.D.; Dahlke, D.B.; Shiono, O.M.; Donadon, A.R.; Camargo, J.C. Power quality analyses of a large scale photovoltaic system. In Proceedings of the 2014 5th International Renewable Energy Congress (IREC), Hammamet, Tunisia, 25–27 March 2014; pp. 1–6. [[CrossRef](#)]
8. Abubakar, A.; Almeida, C.F.M. Analysis of Battery Energy Storage System Sizing in Isolated PV Systems Considering a Novel Methodology and Panel Manufacturers Recommended Methodology. In Proceedings of the 2020 IEEE PES Transmission & Distribution Conference and Exhibition—Latin America (T&D LA), Montevideo, Uruguay, 28 September–2 October 2020; pp. 1–6. [[CrossRef](#)]
9. Dulout, J.; Jammes, B.; Alonso, C.; Anvari-Moghaddam, A.; Luna, A.; Guerrero, J.M. Optimal sizing of a lithium battery energy storage system for grid-connected photovoltaic systems. In Proceedings of the 2017 IEEE Second International Conference on DC Microgrids (ICDCM), Nuremberg, Germany, 27–29 June 2017; pp. 582–587. [[CrossRef](#)]

10. Abubakar, A.; Gemignani, M.; Almeida, C.F.M. Battery Storage System Sizing Using Synthetic Series Data. In Proceedings of the 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC), Chicago, IL, USA, 16–21 June 2019; pp. 1578–1583. [[CrossRef](#)]
11. Feldman, D.; Margolis, R. *Q4 2019/Q1 2020 Solar Industry Update*; National Renewable Energy Laboratory: Denver, CO, USA, 2020.
12. Paul, E.; Bray, D. *Evolution of Solar Operating Practices: Advanced O&M Benefits from Module-Level Monitoring Solution Deployment Brief*; AltaTerra Research: Palo Alto, CA, USA, 2012.
13. Zhao, Y.; Ball, R.; Mosesian, J.; De Palma, J.-F.; Lehman, B. Graph-Based Semi-supervised Learning for Fault Detection and Classification in Solar Photovoltaic Arrays. *IEEE Trans. Power Electron.* **2014**, *30*, 2848–2858. [[CrossRef](#)]
14. Maleki, A.; Nazari, M.A.; Pourfayaz, F. Harmony search optimization for optimum sizing of hybrid solar schemes based on battery storage unit. *Energy Rep.* **2020**, *6*, 102–111. [[CrossRef](#)]
15. Nkambule, M.S.; Hasan, A.N.; Ali, A.; Hong, J.; Geem, Z.W. Comprehensive Evaluation of Machine Learning MPPT Algorithms for a PV System Under Different Weather Conditions. *J. Electr. Eng. Technol.* **2021**, *16*, 411–427. [[CrossRef](#)]
16. Kassim, N.; Sulaiman, S.I.; Othman, Z.; Musirin, I. Harmony search-based optimization of artificial neural network for predicting AC power from a photovoltaic system. In Proceedings of the 2014 IEEE 8th International Power Engineering and Optimization Conference (PEOCO2014), Langkawi, Malaysia, 24–25 March 2014; pp. 504–507. [[CrossRef](#)]
17. Abbas, M.; Zhang, D. A smart fault detection approach for PV modules using Adaptive Neuro-Fuzzy Inference framework. *Energy Rep.* **2021**, *7*, 2962–2975. [[CrossRef](#)]
18. Lu, X.; Lin, P.; Cheng, S.; Fang, G.; He, X.; Chen, Z.; Wu, L. Fault diagnosis model for photovoltaic array using a dual-channels convolutional neural network with a feature selection structure. *Energy Convers. Manag.* **2021**, *248*, 114777. [[CrossRef](#)]
19. Madeti, S.R.; Singh, S.N. Modeling of PV system based on experimental data for fault detection using kNN method. *Sol. Energy* **2018**, *173*, 139–151. [[CrossRef](#)]
20. Chen, S.; Li, X.; Meng, Y.; Xie, Z. Wavelet-based protection strategy for series arc faults interfered by multicomponent noise signals in grid-connected photovoltaic systems. *Sol. Energy* **2019**, *183*, 327–336. [[CrossRef](#)]
21. Hussain, M.; Dhimish, M.; Titarenko, S.; Mather, P. Artificial neural network based photovoltaic fault detection algorithm integrating two bi-directional input parameters. *Renew. Energy* **2020**, *155*, 1272–1292. [[CrossRef](#)]
22. Mellit, A.; Kalogirou, S.A. *A Survey on the Application of Artificial Intelligence Techniques for Photovoltaic Systems*, 3rd ed.; Elsevier Ltd.: Amsterdam, The Netherlands, 2018.
23. Chen, Z.; Yu, H.; Luo, L.; Wu, L.; Zheng, Q.; Wu, Z.; Cheng, S.; Lin, P. Rapid and accurate modeling of PV modules based on extreme learning machine and large datasets of I-V curves. *Appl. Energy* **2021**, *292*, 116929. [[CrossRef](#)]
24. Alhejji, A.; Mosaad, M.I. Performance enhancement of grid-connected PV systems using adaptive reference PI controller. *Ain Shams Eng. J.* **2021**, *12*, 541–554. [[CrossRef](#)]
25. Jung, S.; Jeoung, J.; Kang, H.; Hong, T. Optimal planning of a rooftop PV system using GIS-based reinforcement learning. *Appl. Energy* **2021**, *298*, 117239. [[CrossRef](#)]
26. Vergara, P.P.; Salazar, M.; Giraldo, J.S.; Palensky, P. Optimal dispatch of PV inverters in unbalanced distribution systems using Reinforcement Learning. *Int. J. Electr. Power Energy Syst.* **2022**, *136*, 107628. [[CrossRef](#)]
27. Zhou, Y.; Zheng, S.; Zhang, G. Machine learning-based optimal design of a phase change material integrated renewable system with on-site PV, radiative cooling and hybrid ventilations—study of modelling and application in five climatic regions. *Energy* **2020**, *192*, 116608. [[CrossRef](#)]
28. Singh, Y.; Pal, N. Reinforcement learning with fuzzified reward approach for MPPT control of PV systems. *Sustain. Energy Technol. Assess.* **2021**, *48*, 101665. [[CrossRef](#)]
29. Xie, Z.; Wu, Z. Maximum power point tracking algorithm of PV system based on irradiance estimation and multi-Kernel extreme learning machine. *Sustain. Energy Technol. Assess.* **2021**, *44*, 101090. [[CrossRef](#)]
30. Avila, L.; De Paula, M.; Trimboli, M.; Carlucho, I. Deep reinforcement learning approach for MPPT control of partially shaded PV systems in Smart Grids. *Appl. Soft Comput.* **2020**, *97*, 106711. [[CrossRef](#)]
31. Behera, M.K.; Saikia, L.C. A new combined extreme learning machine variable steepest gradient ascent MPPT for PV system based on optimized PI-FOI cascade controller under uniform and partial shading conditions. *Sustain. Energy Technol. Assess.* **2020**, *42*, 100859. [[CrossRef](#)]
32. Zafar, M.H.; Khan, N.M.; Mirza, A.F.; Mansoor, M.; Akhtar, N.; Qadir, M.U.; Khan, N.A.; Moosavi, S.K.R. A novel meta-heuristic optimization algorithm based MPPT control technique for PV systems under complex partial shading condition. *Sustain. Energy Technol. Assess.* **2021**, *47*, 101367. [[CrossRef](#)]
33. Lin, P.; Peng, Z.; Lai, Y.; Cheng, S.; Chen, Z.; Wu, L. Short-term power prediction for photovoltaic power plants using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets. *Energy Convers. Manag.* **2018**, *177*, 704–717. [[CrossRef](#)]
34. Mishra, M.; Dash, P.B.; Nayak, J.; Naik, B.; Swain, S.K. Deep learning and wavelet transform integrated approach for short-term solar PV power prediction. *J. Int. Meas. Confed.* **2020**, *166*, 108250. [[CrossRef](#)]
35. Kow, K.W.; Wong, Y.W.; Rajkumar, R.; Isa, D. An intelligent real-time power management system with active learning prediction engine for PV grid-tied systems. *J. Clean. Prod.* **2018**, *205*, 252–265. [[CrossRef](#)]
36. Behera, M.K.; Nayak, N. A comparative study on short-term PV power forecasting using decomposition based optimized extreme learning machine algorithm. *Eng. Sci. Technol. Int. J.* **2020**, *23*, 156–167. [[CrossRef](#)]

37. Li, P.; Zhou, K.; Lu, X.; Yang, S. A hybrid deep learning model for short-term PV power forecasting. *Appl. Energy* **2020**, *259*, 114216. [CrossRef]
38. Livera, A.; Florides, M.; Theristis, M.; Makrides, G.; Georghiou, G.E. Failure diagnosis of short- and open-circuit fault conditions in PV systems. In Proceedings of the 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC), Waikoloa, HI, USA, 10–15 June 2018. [CrossRef]
39. Pillai, D.S.; Rajasekar, N. A comprehensive review on protection challenges and fault diagnosis in PV systems. *Renew. Sustain. Energy Rev.* **2018**, *91*, 18–40. [CrossRef]
40. Ghosh, R.; Das, S.; Panizrahi, C.K. Classification of Different Types of Faults in a Photovoltaic System. In Proceedings of the 7th IEEE International Conference on Computation of Power, Energy, Information and Communication, Chennai, India, 28–29 March 2018; pp. 121–128.
41. Nguyen, X.H. Matlab/Simulink Based Modeling to Study Effect of Partial Shadow on Solar Photovoltaic Array. *Environ. Syst. Res.* **2015**, *4*, 20. [CrossRef]
42. Patel, H.; Agarwal, V. MATLAB-Based Modeling to Study the Effects of Partial Shading on PV Array Characteristics. *IEEE Trans. Energy Convers.* **2008**, *23*, 302–310. [CrossRef]
43. Spooner, E.; Wilmot, N. Safety issues, arcing and fusing in PV arrays. In Proceedings of the 3rd International Solar Energy Society Conference—Asia Pacific Region (ISES-AP-08) Incorporating the 46th ANZSES Conference, Sydney, Australia, 25–28 November 2008.
44. Xia, K.; He, Z.; Yuan, Y.; Wang, Y.; Xu, P. An arc fault detection system for the household photovoltaic inverter according to the DC bus currents. In Proceedings of the 2015 18th International Conference on Electrical Machines and Systems (ICEMS), Pattaya City, Thailand, 25–28 October 2015; pp. 1687–1690. [CrossRef]
45. Chen, L.; Li, S.; Wang, X. Quickest Fault Detection in Photovoltaic Systems. *IEEE Trans. Smart Grid* **2018**, *9*, 1835–1847. [CrossRef]
46. Zhao, Y.; Lehman, B.; de Palma, J.-F.; Mosesian, J.; Lyons, R. Challenges to overcurrent protection devices under line-line faults in solar photovoltaic arrays. In Proceedings of the 2011 IEEE Energy Conversion Congress and Exposition, Phoenix, AZ, USA, 17–22 September 2011. [CrossRef]
47. Zhao, Y. Fault Analysis in Solar Photovoltaic Arrays. Ph.D. Thesis, Northeastern University, Boston, MA, USA, 2010.
48. Yi, Z.; Etemadi, A.H. Line-to-Line Fault Detection for Photovoltaic Arrays Based on Multiresolution Signal Decomposition and Two-Stage Support Vector Machine. *IEEE Trans. Ind. Electron.* **2017**, *64*, 8546–8556. [CrossRef]
49. Appiah, A.Y.; Zhang, X.; Ayawli, B.B.K.; Kyeremeh, F. Review and Performance Evaluation of Photovoltaic Array Fault Detection and Diagnosis Techniques. *Int. J. Photoenergy* **2019**, *2019*, 6953530. [CrossRef]
50. Zhao, Y.; De Palma, J.-F.; Mosesian, J.; Lyons, R.; Lehman, B. Line-Line Fault Analysis and Protection Challenges in Solar Photovoltaic Arrays. *IEEE Trans. Ind. Electron.* **2013**, *60*, 3784–3795. [CrossRef]
51. Hua, C.-C.; Ku, P.-K. Implementation of a Stand-Alone Photovoltaic Lighting System with MPPT, Battery Charger and High Brightness LEDs. In Proceedings of the 2005 International Conference on Power Electronics and Drives Systems, Kuala Lumpur, Malaysia, 28 November–1 December 2005; pp. 1601–1605.
52. Wang, Y.; Li, Y.; Ruan, X. High-Accuracy and Fast-Speed MPPT Methods for PV String Under Partially Shaded Conditions. *IEEE Trans. Ind. Electron.* **2016**, *63*, 235–245. [CrossRef]
53. Chan, F.; Calleja, H. Reliability: A New Approach in Design of Inverters for PV Systems. In Proceedings of the 2006 IEEE International Power Electronics Congress, Puebla, Mexico, 16–18 October 2006; pp. 1–6. [CrossRef]
54. Mellit, A.; Tina, G.M.; Kalogirou, S.A. Fault detection and diagnosis methods for photovoltaic systems: A review. *Renew. Sustain. Energy Rev.* **2018**, *91*, 1–17. [CrossRef]
55. Davarifar, M.; Rabhi, A.; El Hajjaji, A. Comprehensive Modulation and Classification of Faults and Analysis Their Effect in DC Side of Photovoltaic System. *Energy Power Eng.* **2013**, *5*, 230–236. [CrossRef]
56. Akram, M.N.; Lotfifard, S. Modeling and Health Monitoring of DC Side of Photovoltaic Array. *IEEE Trans. Sustain. Energy* **2015**, *6*, 1245–1253. [CrossRef]
57. Schimpf, F.; Norum, L.E. Recognition of electric arcing in the DC-wiring of photovoltaic systems. In Proceedings of the Intelec 2009—31st International Telecommunications Energy Conference, Incheon, Korea, 18–22 October 2009; pp. 1–6. [CrossRef]
58. Omer, A. Renewable energy resources for electricity generation in Sudan. *Renew. Sustain. Energy Rev.* **2007**, *11*, 1481–1497. [CrossRef]
59. Russell, S.; Norvig, P. *Artificial Intelligence: A Modern Approach*; Pearson Education: Pearson, London, 2010; Volume 4.
60. Mitchell, T. *Machine Learning*; Springer: New York, NY, USA, 2011; Volume 17.
61. Klass, L. Machine Learning—Definition and Application Examples. Spotlight Metal. 2018. Available online: https://www.spotlightmetal.com/machine-learning--definition-and-application-examples-a-746226/?cmp=go-ta-art-trf-SLM_DSA-20180820&gclid=CjwKCAjwkoz7BRBPEiwAeKw3q30qrjWJ-kiSAkfp6E6Oe_BxzFqk66RL3o2idJPKF1GBXIC94LgOuBoCTwMQAvD_BwE (accessed on 17 September 2020).
62. Simon, H. *Neural Networks and Learning*; Prentice Hall: Hoboken, NJ, USA, 2009; Volume 127.
63. Cormen, T.; Leiserson, C.; Rivest, R.; Stein, C. *Introduction to Algorithms*; MIT Press: Cambridge, MA, USA, 2015; Volume 1.
64. Chang, M.; Hsu, C. PV O&M Optimization by AI practice. In Proceedings of the 36th European Photovoltaic Solar Energy Conference and Exhibition, Marseille, French, 9–13 September 2019. [CrossRef]

65. Samara, S.; Natsheh, E. Intelligent PV Panels Fault Diagnosis Method Based on NARX Network and Linguistic Fuzzy Rule-Based Systems. *Sustainability* **2020**, *12*, 2011. [CrossRef]
66. Chine, W.; Mellit, A.; Lughi, V.; Malek, A.; Sulligoi, G.; Pavan, A.M. A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. *Renew. Energy* **2016**, *90*, 501–512. [CrossRef]
67. Livera, A.; Makrides, G.; Sutterlueti, J.; Georghiou, G.E. Advanced Failure Detection Algorithms and Performance Outlier Decision Classification for Grid-connected PV Systems. In Proceedings of the 33rd EU PVSEC, Amsterdam, The Netherlands, 25–29 September 2017; pp. 1–7. [CrossRef]
68. Chouder, A.; Silvestre, S. Automatic supervision and fault detection of PV systems based on power losses analysis. *Energy Convers. Manag.* **2010**, *51*, 1929–1937. [CrossRef]
69. Syafaruddin; Karatepe, E.; Hiyama, T. Controlling of artificial neural network for fault diagnosis of photovoltaic array. In Proceedings of the 16th International Conference on Intelligent System Applications to Power Systems, Hersonissos, Greece, 25–28 September 2011; pp. 1–6. [CrossRef]
70. Spataru, S.; Sera, D.; Kerekes, T.; Teodorescu, R. Diagnostic method for photovoltaic systems based on light I–V measurements. *Sol. Energy* **2015**, *119*, 29–44. [CrossRef]
71. Zhao, Y.; Liu, Q.; Li, D.; Kang, D.; Lv, Q.; Shang, L. Hierarchical Anomaly Detection and Multimodal Classification in Large-Scale Photovoltaic Systems. *IEEE Trans. Sustain. Energy* **2018**, *10*, 1351–1361. [CrossRef]
72. Mohamed, A.H.; Nassar, A. New Algorithm for Fault Diagnosis of Photovoltaic Energy Systems. *Int. J. Comput. Appl.* **2015**, *114*, 26–31. [CrossRef]
73. Caruana, R. Multitask Learning. *Mach. Learn.* **1997**, *28*, 41–75. [CrossRef]
74. Dey, A. Machine Learning Algorithms: A Review. *Int. J. Comput. Sci. Inf. Technol.* **2016**, *7*, 1174–1179.
75. Opitz, D.W.; Maclin, R. Popular Ensemble Methods: An Empirical Study. *J. Artif. Intell. Res.* **1999**, *11*, 169–198. [CrossRef]
76. Sharma, V.; Rai, S.; Dev, A. A Comprehensive Study of Artificial Neural Networks. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **2012**, *2*, 278–284.
77. Wikipedia. Instance-Based Learning. Available online: https://en.wikipedia.org/wiki/Instance-based_learning (accessed on 31 March 2021).
78. Al-Sahaf, H.; Bi, Y.; Chen, Q.; Lensen, A.; Mei, Y.; Sun, Y.; Tran, B.; Xue, B.; Zhang, M. A survey on evolutionary machine learning. *J. R. Soc. N. Z.* **2019**, *49*, 205–228. [CrossRef]
79. Yar, M.H.; Rahmati, V.; Oskouei, H.R.D. A Survey on Evolutionary Computation: Methods and Their Applications in Engineering. *Mod. Appl. Sci.* **2016**, *10*, 131. [CrossRef]
80. Pavan, A.M.; Mellit, A.; De Pieri, D.; Kalogirou, S. A comparison between BNN and regression polynomial methods for the evaluation of the effect of soiling in large scale photovoltaic plants. *Appl. Energy* **2013**, *108*, 392–401. [CrossRef]
81. Mackay, D.J.C. A Practical Bayesian Framework for Backpropagation Networks. *Neural Comput.* **1992**, *4*, 448–472. [CrossRef]
82. Aziz, F.; Haq, A.U.; Ahmad, S.; Mahmoud, Y.; Jalal, M.; Ali, U. A Novel Convolutional Neural Network-Based Approach for Fault Classification in Photovoltaic Arrays. *IEEE Access* **2020**, *8*, 41889–41904. [CrossRef]
83. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems—Volume 1; Curran Associates Inc.: Red Hook, NY, USA, 2012; pp. 1097–1105.
84. Liu, G.; Yu, W. A fault detection and diagnosis technique for solar system based on Elman neural network. In Proceedings of the 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chengdu, China, 15–17 December 2017; pp. 473–480. [CrossRef]
85. Chao, K.-H.; Chiu, C.-L.; Li, C.-J.; Chang, Y.-C. A novel neural network with simple learning algorithm for islanding phenomenon detection of photovoltaic systems. *Expert Syst. Appl.* **2011**, *38*, 12107–12115. [CrossRef]
86. Garoudja, E.; Chouder, A.; Kara, K.; Silvestre, S. An enhanced machine learning based approach for failures detection and diagnosis of PV systems. *Energy Convers. Manag.* **2017**, *151*, 496–513. [CrossRef]
87. Mekki, H.; Mellit, A.; Salhi, H. Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules. *Simul. Model. Pract. Theory* **2016**, *67*, 1–13. [CrossRef]
88. Coleman, A.; Zalewski, J. Intelligent fault detection and diagnostics in solar plants. In Proceedings of the 6th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems, Prague, Czech Republic, 15–17 September 2011; Volume 2, pp. 948–953. [CrossRef]
89. Nguyen, D.D.; Lehman, B.; Kamarthi, S. Performance evaluation of solar photovoltaic arrays including shadow effects using neural network. In Proceedings of the 2009 IEEE Energy Conversion Congress and Exposition, San Jose, CA, USA, 20–24 September 2009; pp. 3357–3362. [CrossRef]
90. Riley, D.; Johnson, J. Photovoltaic prognostics and health management using learning algorithms. In Proceedings of the 2012 38th IEEE Photovoltaic Specialists Conference, Austin, TX, USA, 3–8 June 2012; pp. 1535–1539. [CrossRef]
91. Jufri, F.H.; Oh, S.; Jung, J. Development of Photovoltaic abnormal condition detection system using combined regression and Support Vector Machine. *Energy* **2019**, *176*, 457–467. [CrossRef]
92. Spataru, S.; Sera, D.; Kerekes, T.; Teodorescu, R. Photovoltaic array condition monitoring based on online regression of performance model. In Proceedings of the 2013 IEEE 39th Photovoltaic Specialists Conference (PVSC), Tampa, FL, USA, 16–21 June 2013; pp. 815–820. [CrossRef]

93. King, D.; Kratochvil, J.; Boyson, W. *Photovoltaic Array Performance Model*; Department of Energy: Albuquerque, NM, USA, 2004; p. 87105-0752.
94. Shapsough, S.; Dhaouadi, R.; Zualkernan, I. Using Linear Regression and Back Propagation Neural Networks to Predict Performance of Soiled PV Modules. *Procedia Comput. Sci.* **2019**, *155*, 463–470. [[CrossRef](#)]
95. Rezgui, W.; Mouss, N.K.; Mouss, L.-H.; Mouss, M.D.; Benbouzid, M. A regression algorithm for the smart prognosis of a reversed polarity fault in a photovoltaic generator. In Proceedings of the 2014 First International Conference on Green Energy ICGE, Sfax, Tunisia, 25–27 March 2014; pp. 134–138. [[CrossRef](#)]
96. Zhao, Y.; Yang, L.; Lehman, B.; de Palma, J.-F.; Mosesian, J.; Lyons, R. Decision tree-based fault detection and classification in solar photovoltaic arrays. In Proceedings of the 2012 Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC), Orlando, FL, USA, 5–9 February 2012; pp. 93–99. [[CrossRef](#)]
97. Benkercha, R.; Moulahoum, S. Fault detection and diagnosis based on C4.5 decision tree algorithm for grid connected PV system. *Sol. Energy* **2018**, *173*, 610–634. [[CrossRef](#)]
98. Rosner, B. American Society for Quality Percentage Points for a Generalized ESD Many-Outlier Procedure. *Technometrics* **1983**, *25*, 165–172. [[CrossRef](#)]
99. Witten, I.; Frank, I.H. *Data Mining—Practical Machine Learning Tools and Techniques with JAVA Implementations*; Morgan Kaufmann: Burlington, MA, USA, 2002; Volume 31.
100. Rezgui, W.; Mouss, L.-H.; Mouss, N.K.; Mouss, M.D.; Benbouzid, M. A smart algorithm for the diagnosis of short-circuit faults in a photovoltaic generator. In Proceedings of the 2014 First International Conference on Green Energy ICGE, Sfax, Tunisia, 25–27 March 2014; pp. 139–143. [[CrossRef](#)]
101. Rezgui, W.; Mouss, K.-N.; Mouss, L.-H.; Mouss, M.D.; Amirat, Y.; Benbouzid, M. Optimization of SVM Classifier by k-NN for the Smart Diagnosis of the Short-Circuit and Impedance Faults in a PV Generator. *Int. Rev. Model. Simul.* **2014**, *7*, 863. [[CrossRef](#)]
102. Hu, T.; Zheng, M.; Tan, J.; Zhu, L.; Miao, W. Intelligent photovoltaic monitoring based on solar irradiance big data and wireless sensor networks. *Ad Hoc Netw.* **2015**, *35*, 127–136. [[CrossRef](#)]
103. Ducange, P.; Fazzolari, M.; Lazzarini, B.; Marcelloni, F. An intelligent system for detecting faults in photovoltaic fields. In Proceedings of the 2011 11th International Conference on Intelligent Systems Design and Applications, Cordoba, Spain, 22–24 November 2011; pp. 1341–1346. [[CrossRef](#)]
104. Takagi, T.; Sugeno, M. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst. Man Cybern.* **1985**, *SMC-15*, 116–132. [[CrossRef](#)]
105. Karakose, M.; Firildak, K. A shadow detection approach based on fuzzy logic using images obtained from PV array. In Proceedings of the 2015 6th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO), Istanbul, Turkey, 27–29 May 2015; pp. 1–5. [[CrossRef](#)]
106. Grichting, B.; Goette, J.; Jacomet, M.; Benjamin, G. Cascaded fuzzy logic based arc fault detection in photovoltaic applications. In Proceedings of the 2015 International Conference on Clean Electrical Power (ICCEP), Taormina, Italy, 16–18 June 2015; pp. 178–183. [[CrossRef](#)]
107. Spataru, S.; Sera, D.; Kerekes, T.; Teodorescu, R. Detection of increased series losses in PV arrays using Fuzzy Inference Systems. In Proceedings of the 2012 38th IEEE Photovoltaic Specialists Conference, Austin, TX, USA, 3–8 June 2012; pp. 464–469. [[CrossRef](#)]
108. Yi, Z.; Etemadi, A.H. Fault Detection for Photovoltaic Systems Based on Multi-Resolution Signal Decomposition and Fuzzy Inference Systems. *IEEE Trans. Smart Grid* **2017**, *8*, 1274–1283. [[CrossRef](#)]
109. Dhimish, M.; Holmes, V.; Mehrdadi, B.; Dales, M. Diagnostic method for photovoltaic systems based on six layer detection algorithm. *Electr. Power Syst. Res.* **2017**, *151*, 26–39. [[CrossRef](#)]
110. Dhimish, M.; Holmes, V.; Mehrdadi, B.; Dales, M.; Mather, P. Photovoltaic fault detection algorithm based on theoretical curves modelling and fuzzy classification system. *Energy* **2017**, *140*, 276–290. [[CrossRef](#)]
111. Yao, X.; Herrera, L.; Ji, S.; Zou, K.; Wang, J. Characteristic Study and Time-Domain Discrete- Wavelet-Transform Based Hybrid Detection of Series DC Arc Faults. *IEEE Trans. Power Electron.* **2014**, *29*, 3103–3115. [[CrossRef](#)]
112. Ray, P.K.; Mohanty, A.; Panigrahi, B.K.; Rout, P.K. Modified wavelet transform based fault analysis in a solar photovoltaic system. *Optik* **2018**, *168*, 754–763. [[CrossRef](#)]
113. Kim, I.-S. On-line fault detection algorithm of a photovoltaic system using wavelet transform. *Sol. Energy* **2016**, *126*, 137–145. [[CrossRef](#)]
114. Davarifar, M.; Rabhi, A.; Hajjaji, A.; Kamal, E.; Daneshifar, Z. Partial shading fault diagnosis in PV system with discrete wavelet transform (DWT). In Proceedings of the 2014 International Conference on Renewable Energy Research and Application (ICRERA), Milwaukee, WI, USA, 19–22 October 2014; pp. 810–814. [[CrossRef](#)]
115. Li, X.; Yang, P.; Ni, J.; Zhao, J. Fault diagnostic method for PV array based on improved wavelet neural network algorithm. In Proceedings of the 11th World Congress on Intelligent Control and Automation, Shenyang, China, 29 June–4 July 2014; pp. 1171–1175. [[CrossRef](#)]
116. Mansouri, M.; Al-Khazraji, A.; Hajji, M.; Harkat, M.F.; Nounou, H.; Nounou, M. Wavelet optimized EWMA for fault detection and application to photovoltaic systems. *Sol. Energy* **2018**, *167*, 125–136. [[CrossRef](#)]
117. Momeni, H.; Sadoogi, N.; Farrokhifar, M.; Gharibeh, H.F. Fault Diagnosis in Photovoltaic Arrays Using GBSSL Method and Proposing a Fault Correction System. *IEEE Trans. Ind. Inform.* **2020**, *16*, 5300–5308. [[CrossRef](#)]

118. Wu, Y.; Chen, Z.; Wu, L.; Lin, P.; Cheng, S.; Lu, P. An Intelligent Fault Diagnosis Approach for PV Array Based on SA-RBF Kernel Extreme Learning Machine. *Energy Procedia* **2017**, *105*, 1070–1076. [[CrossRef](#)]
119. Yagi, Y.; Kishi, H.; Hagihara, R.; Tanaka, T.; Kozuma, S.; Ishida, T.; Waki, M.; Tanaka, M.; Kiyama, S. Diagnostic technology and an expert system for photovoltaic systems using the learning method. *Sol. Energy Mater. Sol. Cells* **2003**, *75*, 655–663. [[CrossRef](#)]
120. Kurokawa, K.; Uchida, D.; Otani, K.; Sugiura, T. Realistic PV performance values obtained by a number of grid-connected systems in Japan. In Proceedings of the NorthSun'99, The 8th International Conference on Solar Energy in High Latitudes, Edmonton, AB, Canada, 11–14 August 1999.
121. Garoudja, E.; Kara, K.; Chouder, A.; Silvestre, S.; Kichou, S. Efficient fault detection and diagnosis procedure for photovoltaic systems. In Proceedings of the 2016 8th International Conference on Modelling, Identification and Control (ICMIC), Algiers, Algeria, 15–17 November 2016; pp. 851–856. [[CrossRef](#)]
122. Chen, Z.; Wu, L.; Cheng, S.; Lin, P.; Wu, Y.; Lin, W. Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I-V characteristics. *Appl. Energy* **2017**, *204*, 912–931. [[CrossRef](#)]
123. Zhao, Y.; Lehman, B.; Ball, R.; Mosesian, J.; de Palma, J.-F. Outlier detection rules for fault detection in solar photovoltaic arrays. In Proceedings of the 2013 Twenty-Eighth Annual IEEE Applied Power Electronics Conference and Exposition (APEC), Long Beach, CA, USA, 17–21 March 2013; pp. 2913–2920. [[CrossRef](#)]
124. Shariff, F.; Rahim, N.A.; Hew, W.P. Zigbee-based data acquisition system for online monitoring of grid-connected photovoltaic system. *Expert Syst. Appl.* **2015**, *42*, 1730–1742. [[CrossRef](#)]
125. Lin, H.; Chen, Z.; Wu, L.; Lin, P.; Cheng, S. On-line Monitoring and Fault Diagnosis of PV Array Based on BP Neural Network Optimized by Genetic Algorithm. *Lect. Notes Comput. Sci.* **2015**, *9426*, 102–112. [[CrossRef](#)]
126. Kuo, C.-L.; Chen, J.-L.; Chen, S.-J.; Kao, C.-C.; Yau, H.-T.; Lin, C.-H. Photovoltaic Energy Conversion System Fault Detection Using Fractional-Order Color Relation Classifier in Microdistribution Systems. *IEEE Trans. Smart Grid* **2015**, *8*, 1163–1172. [[CrossRef](#)]
127. Jones, C.B.; Stein, J.S.; Gonzalez, S.; King, B.H. Photovoltaic system fault detection and diagnostics using Laterally Primed Adaptive Resonance Theory neural network. In Proceedings of the 2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC), New Orleans, LA, USA, 14–19 June 2015; pp. 1–6. [[CrossRef](#)]
128. Khoshnami, A.; Sadeghkhani, I. Sample entropy-based fault detection for photovoltaic arrays. *IET Renew. Power Gener.* **2018**, *12*, 1966–1976. [[CrossRef](#)]