MODELS, DETECTION METHODS, AND CHALLENGES IN DC ARC FAULT: A REVIEW

Alaa Hamza Omran, Dalila Mat Said, Siti Maherah Hussin, Sadiq H. Abdulhussain, Nasarudin Ahmad

Centre of Electrical Energy Systems (CEES), School of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
University of Information Technology and Communications, Iraq
Department of Computer Engineering, University of Baghdad, Baghdad, 10071, Iraq

Abstract

The power generation of solar photovoltaic (PV) technology is being implemented in every nation worldwide due to its environmentally clean characteristics. Therefore, PV technology is significantly growing in present applications and usage of PV power systems. Despite the strength of the PV arrays in power systems, the arrays remain susceptible to certain faults. An effective supply requires economic returns, the security of the equipment and humans, precise fault identification, diagnosis, and interruption tools. Meanwhile, the faults in unidentified arc lead to serious fire hazard to commercial, residential, and utility-scale PV systems. To ensure a secure and dependable distribution of electricity, the detection of such hazard is crucial in the early phases of the distribution. In this paper, a detailed review on modern approaches for the identification of DC arc faults in PV is presented. In addition, a thorough comparison is performed between various DC arc-fault models, characteristics, and approaches used for the identification of the faults.

Keywords: DC Arc fault detection, Fault diagnosis, Photovoltaic systems (PV), Fire hazard photovoltaic, Fault detection method

Abstrak

Penjanaan kuasa melalui teknologi fotovoltaik suria atau solar photovoltaic (PV) kini dilaksanakan di setiap negara di seluruh dunia kerana ciri-cirinya yang tidak mencemarkan alam sekitar. Teknologi PV ini semakin berkembang dari segi perlakuanan terkini dan penggunaan sistem kuasa PV. Walaupun susunan PV, atau PV arrays, mempunyai kekuatannya di dalam sistem kuasa, susunan ini tetap terdedah kepada beberapa kesilapan atau faults. Bekalan yang efektif memerlukan puluhan ekonomi, keselamatan peralatan dan manusia, pengenalpastian kesilapan yang tepat, diagnosis serta perkakas penyelanaan (interruption tools). Sementara itu, kesilapan arka yang tidak berjaya dikejarkan akan membawa kepada risiko kebakan pada sistem PV di fasiliti kediaman, komersial, serta berskala utiliti. Untuk memastikan penghantaran elektrik yang selamat dan boleh diharap, pengesanan bahaya seperti ini amatlah penting, terutamanya di peringkat awal pembahagian tersebut. Dalam penulisan ini, ulasan terperinci yang dilakukan ke atas kaedah moden pengesanan kesilapan arka arus terus atau direct current (DC) dalam PV akan dibentangkan. Tambahan juga, pelbagai model kesilapan-arka DC
1.0 INTRODUCTION

Clean energy resources have become increasingly important in the current years as a result of climate changes and decreasing fossil resources. The PV technology is shown to be the appropriate method to fulfil the need for these resources due to the remarkable features including: the reduction in PV module price [1], low maintenance price, short installation time, and environmental compatibility [2], [3]. As a result, a notable increase in the installed capacity of the global PV power plant was observed from 1.3 GW in 2000 to 177 GW in 2014 [3]. However, an increase in the global in stalled capacity to 310 GW was recorded in 2016 [4]. Provided the increase in the solar power generation worldwide, the residential rooftop solar panels or grid-connected PV generation would have an essential function in supporting the main loads and micro-grids. Furthermore, the enhanced PV systems and DC voltage level contribute to significantly possible emergence of DC arc faults, i.e., utility-scaled PV solar farms normally generate a range of voltage from 600 V and 1000 V, while the normal building PV systems generate voltage ranging from 120 V to 600 V in the USA [5], [6].

The declining condition conductors, connectors, cables, and other system components occur due to long-term ageing and weathering impact. Additionally, the insufficient scheduling of maintenance contributes to the possible increase in the DC arc occurrence in PV systems [7]. Arc faults are normally occurred in the PV systems. High temperature plasma developed by the sustained arc may heavily harm the elements in the system [8]. The defects in the system occurred as a result of fire and arc faults in Bakersfield, USA and Mount Holly, the USA in 2009 and 2011, respectively. This phenomenon became prominent and led to the development and enhancement of the related standards and codes [9], [10]. According to the 2011 National Electrical Code (NEC), all rooftop PV systems of DC operating voltage should amount to over 80 V and be equipped with a series of arc-fault circuit interrupters. Furthermore, the importance constitutes all categories of PV systems with a voltage higher than 80 V in 2014 to manage the fire hazard, which is caused by arc faults [11]. Three surveys were associated with the DC arc faults in the PV systems [12]–[14]. Yao et al. [12] presented a short review of the restricted amount of the approaches of identifying DC arc faults in the PV systems, while Alam et al. [13] performed a thorough survey on approaches of identifying and managing the catastrophic faults, including arc faults, ground faults, and line-faults in the PV systems. [14] presented a comprehensive review on the DC series arc fault with a simply listing and reviewing every method studied in the literature without including the challenges of each group method. However, no substantial investigation was performed on the approaches for the diagnosis of the arc faults for PV systems. Also, none of the surveys made a discussion on the capacities and restrictions of various algorithms of identification.

The main objective of this paper is to discuss the modern approaches of detection for the diagnosis of DC arc faults in the PV systems. Discussion and comparison were performed on the functionalities and restrictions of various DC arc models. For the development of an efficient detection algorithm, the understanding of the features and instruments of the DC arc faults is an essential point.

2.0 PV SYSTEM AND PV GENERATOR MODEL

The setup of PV could be categorised based on the power degrees which are: industrial, buildings, residential, and utility scales. These degrees are also organised based on their link to the utility grid, which consists of stand-alone or grid-connected systems. Based on Figure 1 [15], which presents the system structure, the systems primarily consist of PV modules with a connection to the DC/AC inverter, normally through a junction box. Blocking diodes are normally incorporated in the development of each solar panel. Furthermore, the PV module is the result of the merging of PV cells, which generates electrical power upon exposure to light. The systems also cover the components, which play a role in the production of energy in a PV generator.
Furthermore, every cell series is connected to a bypass diode, which functions to prevent the modules from presenting the same behaviour to the receivers and the heating of cells upon partial illumination. It could be seen from Figure 2 that the I-V and P-V curves of the PV generator are according to the elementary cell modelled by an equal circuit. The phenomenon of power loss is considered in the parallel (Rp) and series resistance (Rs). Notably, this is the most frequently used model in the literature regarding PV monitoring, which is also shown in the EN 50530 standard [16]. The parameters of the model presented the following formula, and various monitoring methods are according to the parameters, including series resistance and shunt resistance measurement. However, several references highlighted in this review paper are dependent on the two diode models, [17] and [18], with one of the diode models described in (1).

\[
I_{pv} = I_{ph} - I_0 \left( \frac{U_{pv} + U_{pv,Rs}}{m UT} - 1 \right) - \frac{U_{pv} + U_{pv,Rs}}{R_{p}} \tag{1}
\]

\[
I_0 = C_0 T_{mod}^3 e^{-\frac{U_{gap}}{BT}} \tag{2}
\]

\[
U_T = \frac{K T_{mod}}{\epsilon_0} \tag{3}
\]

\[
T_{mod} = T + \frac{c}{1000 \frac{W}{m^2}} \cdot G \tag{4}
\]

with \(I_{pv}\): Module current (A), \(I_{ph}\): Diode saturation current (A), \(U_{pv}\): Photocurrent (V), \(U_T\): Temperature voltage (V), \(U_{gap}\): Band gap voltage (V), \(R_s\): Series resistance (\(\Omega\)), \(R_p\): Parallel resistance (\(\Omega\)), \(T\): Ambient temperature (K), \(T_{mod}\): Module temperature (K), \(G\): Irradiance (W/m²), \(C\): Temperature model constant, \(e_0\): Elementary charge, \(k\): Boltzmann constant, \(m\): Diode factor.

The inverter is supplied by an MPPT, which utilised the compatibility between the solar array and the grid. As a result, the system achieves its optimum power, as shown in Figure 3.

### 3.0 CATEGORIES OF DC ARC FAULTS

Fault represents the change in the function of an element from its predicted action [19]. Through an arc fault, a new part in the air is developed. This component may be established due to a disruption in the existing carrying conductors or insulation breakdown, which is parallel to the carrying conductors. It is noteworthy that any category of arc fault would damage the PV system and lead to fire, which may contribute to fire hazards and insulation burn-out when flammable substances are present within the area of the PV plant [20] - [25]. Furthermore, provided that if the DC operating voltage is greater than or equal to 80 V, a series of arc-fault protection tools may be essential for NEC-2011 in the PV system. These tools are known as arc-fault circuit interrupters (AFCIs) [26] [27]. Generally, the installation of AFCI devices and ground-fault detection and interruption (GFDI) is performed within the inverter of the PV system, with its front panel being presented to illustrate the categories of fault taken place. Any arc developed as a result of any disruption in current-carrying conductors (CCCs) mostly lead to abrasion from various sources, rodent damage, connector corrosion, and solder disjoint among others. These elements are known as a series of arc faults, while any arc fault between two adjacent conductors with various potentials is identified as parallel arc fault. The parallel and series arc faults are presented in Figure 4. Majority of the AFCI tools have a role in disrupting the processes of the PV system upon the presence of a
series arc fault, and the AFCIs may not be capable of a good response when the parallel arc fault is present. The series and parallel arc-faults lead to the emergence of high-frequency noise in the DC of the PV string, while the parallel arc fault enhances the reduced voltage/current within the PV array. The difference between the parallel and series arc faults could be performed through observation of the rapid reduction of voltage/current, which is related to the heightened noise in DC. All series arc faults could be distinguished through the disconnection of the PV array from the inverter, which occurs when one of the terminals of every PV string is opened. However, module-level disconnects, short-circuiting of the PV string, or module terminals are essential in several parallel arc faults after the disconnection from the inverter to disconnect the arc.

Figure 4 Examples of different DC arc faults in a PV array

4.0 DC ARC MODELS

Although arc models have been investigated since the past century, the first models emphasise on voltage-current (V-I) equations and Finite Element Analysis (FEA), which identify the arc as a normal phenomenon occurring when a mechanical arc furnace process or circuit breaker is opened, instead of a fault [28 - 29].

An FEA model is developed through the simulation of the overall physical operation in the plasma release. Despite the benefits of the FEA model in the investigation of an arc physical features, it has not been used in the research regarding the impacts of arc fault in a circuit. The V-I equation may be applicable in the examination of the steady-state features of an arc, which generally associates the arc current and voltage with an equation, some coefficients identified by the electrode gap lengths, shapes, and materials among others. Although the equation could also be used to estimate the reactions of the steady-state fault towards a DC system, an essential characteristic of the arc is omitted in the V-I equations (see Figure 3). Specifically, this characteristic refers to a plasma release channel, which is related to the diverse high-frequency elements. These elements rely on random molecular collision-dominated operations.

High-frequency elements have a crucial function in identification. To be specific, these elements allow the identification of the presence of an arc fault as they are monitored either by the current or voltage measurements. The transmission of high-frequency elements in a system leads to the development of sensors in a distant location and possibly decreases the number of required local sensors. Meanwhile, the transmission of the high-frequency element possibly results in cross-talking and unneeded tripping in a neighbouring non-faulted section of the circuit. These issues illustrate the importance of the modelling of high-frequency elements to assist the research on the effects of the elements in a wide range of systems under diverse phenomena.

Three categories of models are generally incorporated, the high-frequency component model, the steady-state model (V-I), and the transient model. Each model demonstrates an important element in the analysis of arc fault reaction and could be utilised individually or simultaneously with another model according to the specified application of it.

A. Steady-state Model (V-I equations)

The DC arc fault is frequently specified as high impedance fault due to the strong resistance shown by the channel, which is unavoidable. Therefore, V-I equations are implemented to illustrate the nonlinear resistive features of the arcing. In general, V-I equations is obtained empirically using the current and curve fitting of the arc voltage with a high amount of testing points.

According to the experimental condition, several V-I equations have been proposed relatively similar forms. A comprehensive review of the proposed equations is presented in [29]. Meanwhile, a summary of several equations and a current modified equation is presented in Table (1), in which L refers to the gap length used to estimate the arc length. On the other hand, A, B, C, and D are the coefficients influenced by experimental conditions [30].

The majority of these equations are proposed for high power applications with a large current and gap lengths. The modified Paukert equation incorporates gap lengths in smaller steps compared to the original Paukert equation, which consists of smaller current and gap length ranges. Following that, the modelling of DC arc faults in the scale is proposed, and these faults could also be found in the modern DC power systems with DC bus voltages of several hundred volts.

Another feature incorporated in the Modified Paukert equation is the nonlinear impact of the gap length L on \( V_{\text{arc}} \), as represented by the term \( b + dL \).
in the denominator. Further information on this derivation i can be found in [31].

| Name                  | Equation                                                                 | Experimental  
|-----------------------|--------------------------------------------------------------------------|---------------
| Ayton                 | $V_{arc} = A + Bt + \frac{C + D}{t_i - t_0}$ condition                  | Carbon electrodes |
| Sternitz              | $V_{arc} = A + \sum_{i=0}^{N} \frac{B}{t_i}$ Carbon and magnetite electrodes |
| Nottingham            | $V_{arc} = \frac{a}{t_i}$ in $n$ is related to electrode material: $0.039$ to $0.39$ |
| Poukert               | $V_{arc} = \frac{a}{t_i}$ $t_i$ is the initiation voltage, $0.039$ to $7.78$ in arc: $0.3$ to $100$ |
| Modified Poukert      | $V_{arc} = \frac{a}{t_i}$ $t_i$ is the initiation voltage, $0.3$ to $100$ |

### B. High-frequency Component Model

The randomness of the arc current, which is represented by the noise generating high-frequency elements, is due to the plasma release in the DC arcing channel. Essentially, the molecule and coulomb collision in the plasma channel is a random process, which results in a Gaussian probability distribution of the arc component of the arc current. Furthermore, a zero-mean Gaussian noise function was used in [32] to illustrate the fluctuated voltage in the arc. Following that, further investigation was performed by [30] regarding the Gaussian distribution of an arc current under different conditions. The arc component of the arc current is compatible with the Gaussian distribution, as shown in equation (5):

$$f_x = \frac{1}{\alpha \sqrt{2\pi}} e^{-\left(\frac{x - \mu}{2\sigma^2}\right)^2}$$

(5)

This study illustrates the feasibility of a Gaussian distribution to describe the current randomness of the arc, which further indicates that this distribution is quantitatively associated with the DC component level of the arc current. The Gaussian distribution fitting of measurement noise may also be separated from the arc generation noise. A different approach was adopted in [33] where a value developed by random was added to the arc resistance equation as shown in equation (6):

$$R_{gap} \approx \frac{V_{dc}}{I_{load}} e^{2\alpha(a-1)\Omega}$$

(6)

And to create jitter ($\alpha$) through a random function as shown in equation (7):

$$q^{k+1} = q^k + \text{rand}()^{10}$$

(7)

A coloured wideband noise was proposed in [34] to model the small-signal behaviour of arcing. While the Gaussian function describes the probability distribution related to the value, the “colour” indicates the signal power distribution over a wide frequency range. It was generally assumed that the random noise of a DC arc was either a pink or Brownian noise [35]. Specifically, the development of the pink noise frequently occurs through the filtering of the Gaussian white noise. As can be seen, the modelling of the high-frequency element of arcing could be performed from the elements of frequency domain distribution and probability amplitude.

### C. Transient Model

An arc transient is identified based on the generation of the arc. The test setup is recommended for UL 1699B, and two electrodes are located at a predetermined gap length, while fine steel wool is placed in a polycarbonate tube. When the circuit is energised, a current will flow through the steel wool, which shortens the two separated electrodes. The current will then melt the steel wool through a similar method of the fuse. Furthermore, the melting of steel wool would lead to ionisation of the air and initiate arcing. In this case, the short arc transient is indicated by the sudden current drop in the load circuit and the voltage increase in the electrodes. The measured arc voltage consists of two parts, namely the anode and cathode voltage, and column voltages. To be specific, the anode and cathode voltage is related to the electromotive force (EMF) and is present in the layer nearest to the electrodes [36]. Notably, it is essential in the initiation and sustainability of the arcing channel. Although the anode and cathode voltage are generally independent of the gap length and external circuit, it is dependent on the electrode material. Meanwhile, the column voltage is a result of the flow of current along the resistive arcing channel, and it also relies on the gap length and current. Therefore, when generated by the steel wool with a fixed gap length, the arc voltage reaches its stability in a short duration. However, the model of this short transient has not been studied in detail. Nevertheless, the di/dt produced during this transient has been widely used as a fault indicator in detecting the series DC arc faults. Another common approach to generate an arc is performed by pulling apart two electrodes, which have initially made a contact. This process is performed until a predetermined gap length is reached or when the arc is naturally extinguished while the circuit is energised. In this case, the voltage across arc still consists of the anode and cathode voltage, and column voltage. The cathode and anode voltage are established in a short duration, and it becomes constant during the entire pulling process. However, the column voltage constantly increases throughout the pulling process due to the increase in the gap length. In this case, the transient process consists of two parts, namely the initial sudden current and voltage step change. Moreover, the quasi-stationary stage takes place when the arcing is established and constantly
5.0 DC ARC FAULTS DETECTION METHODS

Extensive research has been made on the identification of AC arc fault, although the development of DC arc fault is low [37]. As the standard associated with the DC arc fault protection in PV system was released in 2011 (UL 1699B), efficacious identification algorithms for DC arc fault and products has become increasingly important [38, 39]. Most of the identification approaches involve the use of measured system voltage data and current, while other approaches incorporate other physical features of arcs. Other than the identification approaches for PV, reviews will be made on several effective identifications approaches for other DC system, including electric vehicle and DC microgrid. It is noteworthy that these approaches are applicable in the PV systems with or without minor changes involved [14]. The detection process of DC series arc fault can be classified according to the utilized method.

A. Model-based Approaches

The model-based approach was suggested in [40] and [41] for the development and verification of the arc fault detection algorithm. The small-signal analysis could be performed on the propagation behaviour of high-frequency elements to assist the creator in identifying the most ideal identification frequency band. Meanwhile, the large-signal analysis could be conducted on the quasi-stationary actions for the assessment of the boundary states for arc faults. This analysis highlights that the arc and system fault modelling is important due to its higher testability of various operating and weather conditions and system topologies compared to the actual condition. Furthermore, the calculation of the difference between the simulated and actual signals could be performed after the establishment of a precise system model. This calculation contributes to the efficient identification of arc faults. The effectiveness of these approaches is mostly restricted by the modelling accuracy. It is also noteworthy that model-based approaches are highly compatible with AI-based approaches (data-driven techniques) due to the high amount of data generated for training purpose [42]. Parameter estimation and detection of PV cell and module are crucial to achieving an accurate model of the PV systems [43]. The production of static and dynamic parameters is reviewed in [44] and [45], respectively. A summary of these methods can be seen in Table 2.

B. Artificial Intelligence-based Approaches

Artificial intelligence is applied in automatic fault detection [46], which involves an artificial network (ANN) and conventional analytical approaches. A two-layer ANN is implemented for the estimation of power according to the temperature and irradiance of the module. Following that, the comparison is made between the estimated power and the actual measurement of it for the identification of fault category and occurrence. Afterwards, the measured open-circuit voltage and value of short circuit value are then compared using the analytical equation, with the calculated values being incorporated. Moreover, this equation consists of a compact structure, high accuracy, and swift identification reaction. It also does not require complex tools or system comprehension. Another method is based on the Artificial Neural Network-Based Model (ANNBM), which functions in identifying loss in the PV panel due to partial shading, as shown in [47]. A Multilayer Perceptron is incorporated for the estimation of electrical output according to environmental data. The monitoring system enables fast identification of system faults through the calculation and analysis of residual faults between the estimated and calculated functions of the ANNBM. Notably, in the ANNBM method, any complex system or mathematical model is not required for the calculation of output power. A proper monitoring method involving probabilistic neural network was highlighted in [48] for the identification and categorisation of faults, including the detection of faulty string in PV array. Accordingly, the suggested method incorporates electrical and environmental data. To validate this method, a new approach to model the PV system was incorporated into the manufacturer datasheet data. It is illustrated in [49] that the hybrid model for the monitoring and estimation of PV module maximum power output considers the nonlinear characteristics of the system, which is known as estimation model (EM). This approach aims towards achieving a more effective estimation method of the actual action of the system. Therefore, correction of the estimation of one-diode model power involves a supervised adaptive resonance theory neural network, while the implementation of the monitoring system is performed through Matlab, followed by its assessment on the sample data. Notably, the suggested model facilitates any PV module technology due to its assessment on the input and output relations. Provided that the simple layer ANN is less compatible in providing a precise solution for the faults in PV modules, another fault identification approach, particularly for short circuit defaults, involves three-layer artificial neural network [50]. Through this approach, the short-circuited PV module is localised in a string. In localising the fault further between strings, addition in the amount of data training and update of the control rule is essential in the system. Furthermore, although the suggested
method has simplicity and accuracy, the implementation of it may be complex in large-scale PV systems. To illustrate, a large memory is required in this implementation as a high number of ANN classifiers are involved. Health management (PHM) system and PV prognostic are illustrated in [51]. This model targets the evaluation of degradation, PV health monitoring system, and further details of the maintenance plans. This approach is according to a duplicate neural network model, including the power input and environmental information. Therefore, the users would have an awareness regarding the decrease in performance. The model also predicts the output of the system according to meteorological parameters. It could be said from the contrast between the estimated and measured AC power that identification is performed failures, which lead to power loss. In the modelling of output power in the ANN health monitoring system, the data regarding the elements or system topology is not essential. The system is also capable of indicating serious failures when monitoring the degradation rate. Finally, a PHM system is capable of estimating the faults in PV according to fault precursors. Although the function of ANN is to estimate the power output from the PV array according to the monitored irradiance in comparison to the calculated power, it is incapable of detecting the factors of the faults. It was stated in [52] that the use of the second artificial intelligence approach, which is known as Bayesian Belief Networks (BBN). The method implemented by the author was similar to the approach shown in [53] for the detection of faults sources. In this research, BBN was developed to indicate the reliance of the measurement variables, which are the representations of particular devices. The measurements achieved by BBN results in an automatic adoption of a decision related to the possible factors of a fault. Introduction of the algorithms according to the fuzzy logic was made in [54 – 56]. Specifically, [54] algorithm was introduced by authors according to the fuzzy mathematics and evidence theories. The suggested study developed new and precise framework of fault identification structure, which was followed by the consolidation of the contrasts between the estimated and calculated value according to the data fusion approach and fuzzy mathematics. It was indicated from the investigations that data fusion is capable of managing the uncertainties from the interaction between PV arrays and localising faults more effectively for large scale PV arrays. An approach based on the measurement of PV module parameters in various situations of operation through the Neuro-Fuzzy method was suggested in [56]. Furthermore, the PV system status is identified by the norms of assessment and comparison. This intelligent system emerged in the Matlab and Simulink environment, and it was found that the diagnosis system was capable of differentiating between a normal and defective operation through the similar defective appearance of interference and noises. Another expert system for PV monitoring was suggested in [57], and it investigates the system database and identifies the energy loss resulting from the inverter defaults. This approach comprises the differentiation between the stored value and real-time output in the database in a fault-free functioning. Notably, a monthly update should be performed on this approach, and it is not applicable for real-time monitoring. As a result, the reduction in power is not identified in the system with the presence of shading. This technique is also according to the shading approach, and it also considers particular states of the monitored system. Moreover, it performs simplification of the processes and preservation of PV plants despite its requirement for various measurement sensors. It involves the standards of the database, performs analysis on the information acquired from the PV system, and affirms if the plant is functioning properly. Two categories of faults were detected in this approach, namely shading and inverter failures. The authors in [58] used the extension theory for PV systems to highlight the two-stage diagnosis system. Through this approach, the identification of diagnosis failures is possible. The primary approach is the detection of fault array in the daytime and the use of light scanning method for the detection of a faulty module in the night. In [59] SVM method was used to detect the series arc fault in the PV which is better than the ANN methods. Table 3 illustrates a brief description of these methods.

C. Statistics-based Methods

Various statistical features could be applied in the identification of DC arc faults, including entropy, RMS value, standard deviation, mean, and the highest value of the input signal. Basing on time-domain analysis, this approach is present in [60–68]. Although the rate of changes in the loop current in the time domain was suggested in [63] to indicate the arc fault event, it would receive the impact of random spike disturbance. It is shown in [61] and [64] that the contrast between the highest and lowest value of current within a certain duration is identified as an indicator. Despite the simplicity of the approach, it is relatively efficacious in identifying arc fault, particularly in the initial phase. Although it is capable of eliminating random interference from the noise, the function of possibly receives an impact from other elements, including the MPPT operation from the inverter during the swift changes in the irradiation level. It is shown in [65, 66] that through the simulation, an outlier analysis-based detection approach obtained 98% of the identification degree of PV series arc fault at 0.01% false alarm rate in a single PV module. Furthermore, the Minimum Covariance Determinant (MCD) predictor is applied to reach the optimum function of the algorithm. It could be seen from the outlier analysis that the current of various PV modules and operating voltage at a similar time instant are placed in the MCD estimator. This was followed by the calculation of the distance of the I-V
characteristic curve between each PV module and the centre of the administration of PV module by the MCD estimator, which could be applied for identification. A Finite Impulse Response (FIR) estimator was applied in [67] to measure the difference of the input voltage signal under the 50 kHz sampling frequency. Initially, the input signal would flow through a band-pass filter with 1 kHz and 7.5 kHz cut-off frequency before being placed into the estimator. Comparison between the current and previous values would be conducted by the estimator. The variance should amount to 0 with an ideal estimation. As the variance goes beyond a predefined threshold value, the detection of an arc fault event is possible. Notably, the implementation of the proposed algorithm is more convenient and less costly although it leads to lower accuracy in comparison to other algorithms. Various detection features were used in [68] for joint detection of arc faults in PV systems. Meanwhile, the calculation of the statistical characteristics, such as the mean and variance of the loop current, is performed in the time domain. In the frequency domain, the calculation is performed on the ratio of frequency contents of the loop current in 1 Hz – 4 kHz to DC and AC components. The identification of an arc event is performed after a minimum of the one-time domain and frequency domain characteristics go beyond the pre-defined threshold value. It is noteworthy that multiple detecting criteria lead to significant improvement in the precision of the algorithm. Currently, an introduction was made in [62] on a statistical detection method according to the arc current entropy. This approach is capable of an effective differentiation of the arc faults from the normal events (non-arc states), including MPPT operation and switch of the inverter. This differentiation is performed by measuring the adjusted Tsallis entropy of loop current for two times. Through Tsallis entropy, the extent of disorder and signal intrinsic behaviour are presented:

\[
M = \sum_{k=1}^{N} p(x_k)^q = 1 - (1 - p(x_k)^q)\frac{E_{Tsallis|M}}{E_{Tsallis}}
\]

\[
p(x = x_k|t = kT) = p(x_k) = \frac{|x_k|^2}{\sum_{i=1}^{N} |x_i|^2}
\]

\[
M' = \sum_{k=1}^{N} p(x_k|M)^q = 1 - (1 - p(x_k|M)^q)\frac{E_{Tsallis|M}}{E_{Tsallis}}
\]

\[
p(x = x_k|M|t = kT) = p(x_k|M) = \frac{|x_k|^2}{\sum_{i=1}^{N} |x_k|^2}
\]

### Table 2 DC Arc Fault Detection Model-based Approaches

<table>
<thead>
<tr>
<th>Ref. No.</th>
<th>Experimentally verified?</th>
<th>Product or Microcontroller</th>
<th>Frequency of Sampling</th>
<th>Accuracy of Test</th>
<th>Time of Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>[41]</td>
<td>No</td>
<td>Not Stated</td>
<td>Not Stated</td>
<td>Not Stated</td>
<td>Not Stated</td>
</tr>
<tr>
<td>[42]</td>
<td>No</td>
<td>Not Specified</td>
<td>Not Stated</td>
<td>Not Stated</td>
<td>Not Stated</td>
</tr>
</tbody>
</table>

### Table 3 DC Arc Fault Detection based Artificial Intelligence

<table>
<thead>
<tr>
<th>Ref. No.</th>
<th>Method</th>
<th>Experiment verified?</th>
<th>Product or Microcontroller</th>
<th>Frequency of Sampling</th>
<th>Accuracy of Test</th>
<th>Time of Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>[46]</td>
<td>ANN and the conventional analytical method</td>
<td>No</td>
<td>Not performed</td>
<td>20 kHz</td>
<td>Not stated</td>
<td>1 ms</td>
</tr>
<tr>
<td>[47]</td>
<td>Artificial neural network-based model (ANNBM)</td>
<td>No</td>
<td>Not performed</td>
<td>1 MHz</td>
<td>Not stated</td>
<td>Not stated</td>
</tr>
<tr>
<td>[48]</td>
<td>Probabilistic neural network (PNN) supervised adaptive resonance theory neural network</td>
<td>No</td>
<td>Not performed</td>
<td>Above 60 MHz</td>
<td>around 98%</td>
<td>Not stated</td>
</tr>
<tr>
<td>[49]</td>
<td>Yes</td>
<td>Not performed</td>
<td>22 kHz</td>
<td>Not stated</td>
<td>Not stated</td>
<td></td>
</tr>
<tr>
<td>[52][53]</td>
<td>Bayesian belief network (BBN)</td>
<td>Yes</td>
<td>FINEMAN</td>
<td>500 MHz</td>
<td>Not stated</td>
<td>Not Stated</td>
</tr>
<tr>
<td>[54]</td>
<td>Yes</td>
<td>Not performed</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated</td>
<td>Not stated</td>
</tr>
<tr>
<td>[59]</td>
<td>SVM</td>
<td>Yes</td>
<td>STM8S103</td>
<td>Above 60 MHz</td>
<td>99.25%</td>
<td>84 ms</td>
</tr>
</tbody>
</table>
Where M represents the modified Tsallis entropy, while q, q’> 0, and x, k refer to the specimens of the interest of the signal, K denotes the size of the sliding window in the initial measurement, and K’ represents the size of sliding window in the measurement which follows it. In every sliding window, p(x_k) and p(x, k, M) amount to 1. The acquired current represents the input signal in the initial phase of the adjusted Tsallis entropy evaluation (\(M\)). As a result of the fast movement and mechanical vibration by the wind, the MPPT algorithm would present variance in M values with particular patterns. Using the appropriate value of K and K’, the interferences could be managed in the second phase of adjusted Tsallis entropy evaluation (\(M’\)), in which M refers to the input signal. Furthermore, a first order infinite impulse response (IIR) filter would be transferred by the \(M’\) with 16 Hz cut-off frequency for the elimination of the DC off-set of M to obtain the detection characteristic \(M_{det}\). The calculation of the threshold value can be performed based on the standard deviation of \(M_{det}\), and the update of it will be made in every 0.5 seconds to be in line with the evolving operating state of the system. The sampling frequency was found to be highly profitable as it amounted to 10 kHz, while the computational load only amounted to 24 N fops in comparison to 5N log, N fops of FFT per sliding window. However, the effectiveness of this approach would be lower when the noise in the environment is higher, and the effects of the impulse reaction of digital IIR filter in the algorithm must be considered. Although the lower sampling frequency and computational attempt are mostly required in the statistics-based fault diagnosis approaches, the effectiveness of these approaches receives a significant impact from the degree of noise from the environment. Table 4 presents a summary of these methods.

<table>
<thead>
<tr>
<th>Ref. No.</th>
<th>Method</th>
<th>Experiment verified?</th>
<th>Product or Microcontroller</th>
<th>Frequency of Sampling</th>
<th>Accuracy of Test</th>
<th>Time of Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>[60]</td>
<td>Statistics</td>
<td>Yes</td>
<td>TI TMS320F28335</td>
<td>10 kHz</td>
<td>Pass all test cases under experimental condition</td>
<td>0.511 s</td>
</tr>
<tr>
<td>[65][66]</td>
<td>Statistics</td>
<td>Yes</td>
<td>Not stated</td>
<td>Not stated</td>
<td>98%</td>
<td>Not stated</td>
</tr>
<tr>
<td>[67]</td>
<td>Statistics</td>
<td>Yes</td>
<td>TI TMS320F2808</td>
<td>50 kHz</td>
<td>Not stated</td>
<td>Less than 175 us</td>
</tr>
<tr>
<td>[68]</td>
<td>Statistics and FFT</td>
<td>Yes</td>
<td>Not performed</td>
<td>more than 8 kHz</td>
<td>Not stated</td>
<td>Not stated</td>
</tr>
</tbody>
</table>

D. Wavelet Transform

Wavelet transform (WT) is a multi-resolution analysis which is capable of the decomposition of the aimed signal into various bands of frequency. In comparison to STFT, adjustability is present in the windowing function of WT. Specifically, larger window size improves the time resolution for low-frequency elements, while smaller window size improves the frequency resolution for high-frequency elements [69]. Wavelet Packet Decomposition (WPD) and Discrete Wavelet Transform (DWT) have been applied for the identification and analysis of arc faults in PV systems [70, 71, 61, 72, 73 – 79]. It could be seen in [69] and [74] that DWT has been applied for the analysis of the arc flash signal. In 7.8 – 62.5 kHz frequency band under 1 MHz sampling rate, DWT is capable of effectively locking the arc instant and identifying the arc fault. The selection of Daubechies (db) wavelets is due to its effectiveness in identifying the discontinuity in the waveforms [76]. It was also recorded that db9 was capable of achieving a positive agreement between computational precision and complexity. Furthermore, 4-level DWT has been applied as an indicator [77], while a real-time arc fault identification approach for PV systems according to 1-level DWT was proposed in [78]. The ratio of the average power of the first level DWT coefficients (50 – 100 kHz) under 200 kHz sampling frequency within a frame to the reference average power was used as the characteristics. An arcing event is established when the ratio is higher than the ratio set for two consecutive frames. The frame size of 128 was selected to reduce the increase in the noise and maintain the precision in localising the time for an arc fault. The calculation of the reference average power was performed based on the first 32 frames of information, which was presumed to be the non-arc state. However, the weakness of it is the failure of it upon the occurrence of arc fault in the algorithm initialisation phase. Based on [79], the implementation of 4-level DWT has been made to obtain the energy information and modulus maximum of 126 – 250 kHz frequency band from the loop current lower than 2000 kHz sampling frequency to identify the parallel arc faults in PV systems. The
investigations found that the two features exhibited 1.85% and 1.47% rejection rates, respectively, with 0% malfunction rate. To achieve further improvement in the effectiveness of the detection, the highest value of the reverse current was represented as an adequate state due to its significant malfunction rate (13.42%) with 0% rejection rate. The three characteristics would develop a new combined criterion, which leads to a notable reduction of the malfunction (0%) and rejection rates (0.875%). Although the high sampling degree is the weakness of the suggested approach, mixing criteria significantly compensate for the insufficiency in the single criterion. The implementation of WPD was made in [73] to obtain the energy of various subbands under 100 kHz sampling frequency. After six levels of decomposition, the detection of series and parallel arc faults is performed when the ratio of the overall square of reconstruction coefficient, which progresses from the highest (781.25 Hz – 50 kHz) to the lowest frequency band (0 – 781.25 Hz), goes beyond the established threshold value for some consecutive analysis durations. In [61] and [75], two-level WPD was implemented to achieve normalisation under 200 kHz sampling frequency. Normalisation was performed on the RMS value of coefficient frequency band of 0 – 25 kHz (including DC off-set) through the frequency band of 25 – 50 kHz within a particular duration (e.g., 10ms). Following that was an increase in the RMS value from 6% to 15%, indicating that the signature was caused by the increase in the series arc fault. The primary contrast between DWT and WPD is illustrated in Figure 6. Besides the decomposition of the lower frequency band at each degree, the decomposition of WPD also occurs at a higher frequency band of the signal. Therefore, more information was offered by WPD compared to DWT, which nearly resulted in a twofold computational burden. More details on these methods presented in Table 5.

E. FFT (Fast Fourier Transform)

FFT (Fast Fourier Transform) is a time-frequency analysis technique used to save DFT computing time [80]. The detection approaches according to the Fast Fourier Transform (FFT) are present in [39, 81-88,72]. Most of the approaches make a comparison between the arc state and non-arc state in terms of the power of the frequency spectrum and amplitude. Based on [39] and [82], the power spectrum of contents in 40 – 100 kHz frequency band was used for the detection of both parallel and series arcs under 250 kHz sampling frequency. The frequency bands of 40 – 80 kHz and 30 – 100 kHz were also incorporated in the 200 kHz sampling frequency in [83] and [84], respectively. However, these approaches might meet their failures upon the presence of electromagnetic interference [87, 88]. In [85], rather than the comparison of the overall spectrum, the frequency spectrum was separated into various sub-bands, which contributed to a significant increase in the accuracy.

Based on [89] and [90], STFT was implemented in the calculation of one of two features for the detection of series arc fault in PV systems under 200 kHz sampling frequency. It was also recorded that the maximum extent of the frequency element, which amounted to over 50 kHz, was not sufficient to indicate the contrast between the arc-state and non-arc-state. Following that was the total frequency contents, which were lower than 50 kHz with the rise of the value within each duration. Continuous supervision was made on the value to detect faults. It was also recorded that STFT was capable of achieving a proper computational load through proper selection of the window size. These methods are illustrated in more details in Table 6.

Furthermore, the detection method that are classified and discussed in detail as mention in this section according to the utilized method, Table 7 shows a brief comparison among these methods according to several factors. The challenges of the methods are depicted in the Challenges Section.

6.0 LIMITATIONS OF EXISTING METHODS

There are many challenges and unresolved issues can affect the performance of the detection process. These issues are listed and discussed according to their affection in each group of methods as follows:

1. Time Domain Based: these methods utilize the information of voltage or current in the time-domain; they can quickly detect arc faults using straightforward DC circuit. Nevertheless, they are not appropriate for AC arc detection, since the changes affection of the normal load current and the noise signals could usually send false trigger for faults detection. Therefore, the time domain methods can be considered as unreliable detection methods.

2. Frequency Domain Based: in these methods, the ejection of a suitable frequency band without noise signals of the switching process play an essential role of arc fault accuracy detection. Therefore, they are not suitable to successfully operate in the common DC arc that includes the noise signals of the switching process as a results of electronics devices multiple power.

3. Modelling Based: in these methods, the high precision of the arc fault detection is significantly depending on the properly modeled of the DC circuit component. However, it is quite difficult to model the component of DC circuit with the required level of the modeling precision.

4. Time-Frequency Based: these methods are directly affected with the presence of the electromagnetic interference.
5. Statistical based: the methods which based this kind of detection are frequently need less sampling with less effort of computational. On the other hand, their efficiency is highly affected by the noise degree at the surrounding environment.

6. Artificial Intelligence: these methods may suffer different challenges as the degradation of the performance that can occurs in many artificial intelligence algorithms. The over-fitting and under-fitting problems which affect the accuracy of the detection methods. And the high computational complexity

<table>
<thead>
<tr>
<th>Ref. No.</th>
<th>Method</th>
<th>Experiment verified?</th>
<th>Product or Microcontroller</th>
<th>Frequency of Sampling</th>
<th>Accuracy of Test</th>
<th>Time of Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>[69]</td>
<td>WPD</td>
<td>Yes</td>
<td>Not performed</td>
<td>100 kHz</td>
<td>100%</td>
<td>0.1 s</td>
</tr>
<tr>
<td>[61],[75]</td>
<td>WPD and Statistics</td>
<td>Yes</td>
<td>TI TMS320F28335</td>
<td>200 kHz</td>
<td>100% at voltage below 300 V and current below 25 A; 60% at 240 V/25 A; 40% at 300 V/25 A</td>
<td>Less than 0.9 s</td>
</tr>
<tr>
<td>[70]</td>
<td>DWT</td>
<td>Yes</td>
<td>Not performed</td>
<td>1 MHz</td>
<td>Not stated</td>
<td>Not stated</td>
</tr>
<tr>
<td>[78]</td>
<td>DWT</td>
<td>Yes</td>
<td>TI RD-195</td>
<td>200 kHz</td>
<td>Not stated</td>
<td>Less than 0.2 s</td>
</tr>
<tr>
<td>[79]</td>
<td>DWT and Statistics</td>
<td>Yes</td>
<td>Not performed</td>
<td>2 MHz</td>
<td>0% malfunction rate and 0.875%</td>
<td>Not stated</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ref. No.</th>
<th>Method</th>
<th>Experiment verified?</th>
<th>Product or Microcontroller</th>
<th>Frequency of Sampling</th>
<th>Accuracy of Test</th>
<th>Time of Detection</th>
</tr>
</thead>
<tbody>
<tr>
<td>[39,82]</td>
<td>FFT</td>
<td>Yes</td>
<td>TI RD-195</td>
<td>250 kHz</td>
<td>Not stated</td>
<td>under 0.2 s</td>
</tr>
<tr>
<td>[83]</td>
<td>FFT</td>
<td>Yes</td>
<td>Not stated</td>
<td>200 kHz</td>
<td>Not stated</td>
<td>under 0.2 s</td>
</tr>
<tr>
<td>[84]</td>
<td>FFT</td>
<td>Yes</td>
<td>Not stated</td>
<td>200 kHz</td>
<td>Not stated</td>
<td>Not stated</td>
</tr>
<tr>
<td>[85]</td>
<td>FFT</td>
<td>Yes</td>
<td>Not performed</td>
<td>250 kHz</td>
<td>Not stated</td>
<td>Not stated</td>
</tr>
<tr>
<td>[86]</td>
<td>FFT</td>
<td>Yes</td>
<td>TI TMS 320F28377D</td>
<td>250 kHz</td>
<td>Pass all test cases under experimental</td>
<td>under 16 ms</td>
</tr>
<tr>
<td>[89,90]</td>
<td>STFT</td>
<td>Yes</td>
<td>Not stated</td>
<td>200 kHz</td>
<td>98.24%</td>
<td>0.5 s</td>
</tr>
</tbody>
</table>
7.0 CONCLUSION

The review has been made on the current formulations of DC arc-fault identification in the literature and patents. Following the increase in the DC system scale and complexity, more sophisticated detection algorithms were proposed. It was concluded in this research that the improved versions of detection algorithms are yet to be developed. Their algorithms would play a role in the management of more complicated environments consisting of spurious noise due to normal operating conditions. Provided that it may be challenging for a single detection strategy to function properly under all conditions, the understanding of the limitations of this strategy in different applications is important. Among the approaches highlighted in the previous segment of this article, the majority of the fault signatures were obtained from the 1 – 100 kHz frequency band. As a result, the most minor effect was developed on the arc fault diagnosis by most of the interferences, with the inverter/converter noise being the exception. Subsequently, omitting the impact of power electronics noise could lead to a notable rise in the precision and dependability of the identification algorithms. Currently, the introduction has been made on various approaches to identification with proper switching noise immunity. Overall, this article has presented a review on DC arc-fault modelling approaches, including the recently formulated detection approaches, which could be applied in PV systems. Moreover, a thorough discussion has also been made on the merits and demerits of various detection approaches, followed by a comparison between them. However, the improved arc fault detection approaches for PV systems with proper self-adaptability, strength and profitability are yet to be identified. For that reason, further attempts are necessary to achieve improvement in the precision and dependability of fault identification.

Acknowledgement

The authors would like to express the appreciation to the Ministry of Higher Education Malaysia (MOHE), the support of the sponsors [Vot Number = Q.J130000.3551.07G53 and R.J130000.4J347] and also to the Universiti Teknologi Malaysia (UTM) for providing the best education and research facilities.
to achieve the aims and goals in research studies and works.

References


