electronics

Article

PV Module Fault Detection Using Combined Artificial Neural Network and Sugeno Fuzzy Logic

Romênia G. Vieira [1], Mahmoud Dhimish [2*], Fábio M. U. de Araújo [3] and Maria I. S. Guerra [1]

[1] Department of Engineering and Technology, Semi-Arid Federal University, Francisco Mota Av., Mossoro, 59625-900, Brazil; romenia.vieira@ufersa.edu.br (R.G.V); izabel.guerra@ufersa.edu.br (M.I.S.G)
[2] Department of Engineering and Technology, University of Huddersfield, Huddersfield, HD1 3DH, United Kingdom
[3] Department of Computer and Automation Engineering, Federal University of Rio Grande do Norte, Natal, 59078-970, Brazil; meneghet@dca.ufrn.br
* Correspondence: M.A.Dhimish@hud.ac.uk

Received: date; Accepted: date; Published: date

Abstract: This work introduces a new fault detection method for photovoltaic systems. The method identifies short-circuited modules and disconnected strings on photovoltaic systems combining two machine learning techniques. The first algorithm is a multilayer feedforward neural network, which uses irradiance, ambient temperature, and power at the maximum power point as input variables. The neural network output enters a Sugeno type fuzzy logic system that precisely determines how many faulty modules are occurring on the power plant. The proposed method was trained using a simulated dataset and is validated using experimental data. The obtained results showed 99.28% accuracy on detecting short-circuited photovoltaic modules and 99.43% on detecting disconnected strings.

Keywords: Photovoltaic system; Photovoltaic faults; Fault detection; ANN networks; Fuzzy logic system.

1. Introduction

Photovoltaic (PV) solar energy has been showing worldwide expansion, reaching an installed capacity of 627 GW [1]. Following this growth, it is essential to ensure the security and reliability of solar power plants. In this perspective, some challenging issues are associated with it, such as faults occurring on PV systems, that may impact the secure operation and the optimal energy harvesting.

The reliability of the PV system can be affected by several factors, such as weather conditions, partial shading, dust/snow accumulation on the modules, wiring losses, aging or malfunctioning of any system component [2]. Some faults could remain undetected by the operators for long periods, and it has the potential to reduce 18.9% of power production [3]. Therefore, it is essential to develop methods capable of detecting and diagnosis faults occurrence in PV systems.

Faults in PV systems can arise on the DC (Direct Current) or the AC (Alternate Current) side. It can affect the PV modules, converters, MPPT (Maximum Power Point Tracking), and storage system on the DC side. PV modules faults are crucial since it is the generation unit of a PV system. Faults occurring on this device could significantly affect the output power. Besides, it could have destructive effects on their efficiency and lifetime [2].

There are various PV modules faults sources, such as mismatch, bypass diode, circuit faults, asymmetrical faults, arc faults, ground faults, and lightning. It can be temporary or permanent, depending on the cause and the period that affects the PV systems performance [4]. The circuit faults, which are the subject of this research, can be open-circuit or short-circuit. In both situations, the PV module does not contribute to power generation, so it could significantly decrease the system performance and remain undetected.

Therefore, some monitoring and diagnosis techniques have been developed in recent years. Diagnosis methods can be classified into two general groups: visual and electric. Visual methods require constant verification by an operator and some specific types of equipment. On the other hand, electrical methods use output signals such as voltage, current, and power to identify faults occurrence [5]. Besides, it is possible to use existing sensors and pieces of equipment, making it more viable for implementing the fault detection method.

In this context, researchers have been exploring different techniques for detecting and diagnosing faults on PV systems. Methods which applies machine learning are widely explored since it offers an alternative way for approaching complex issues. Some of these methods explore detecting the fault in advance, predictively, avoiding massive power losses and damages on PV systems [6]. However, the most common methods search for fault diagnosis in real-time.

Syafaruddin *et al*. [7] developed a diagnosis method using an artificial neural network (ANN). In this case, one neural net was trained for each PV module in order to identify and locate short-circuited modules. The inputs variables are module temperature, irradiance, and maximum power point voltage and current. The method showed promising results, although it was tested only for six modules PV system.

Li *et al.* [8] also applied ANN for fault detection, using the same inputs as [7]. For training, the dataset used was extracted from simulations using MATLAB/Simulink®. The faults detected by the method are degradation, short-circuited modules, and partial shading. However, the method was not experimentally tested. Jiang and Maskell [9] developed a technique for detection combining ANN and analytical method. The ANN forecasts the maximum power point (MPP), using module temperature and irradiance as inputs. The algorithm identifies the fault comparing the provided MPP to the one measured. The faults approached on this method are open-circuited string/module, short-circuited module, partial shading, and malfunctioning MPPT. However, the authors did not test the method experimentally.

Akram and Lotfifard [10] trained a PNN (Probabilistic Neural Network) for detecting short-circuited and open-circuited modules on PV systems. The dataset for training the neural network was assembled by simulation using MATLAB/Simulink® software, and its test showed a maximum error of 3.5%. Nevertheless, it was not tested experimentally, only with simulation data. Further, Garoudja *et al*. [11] also developed a fault detection method using a PNN. Firstly, the PV modules parameters are extracted in order to simulate the studied PV system. The simulation is performed using MATLAB/Simulink® and PSIM® and validated with experimental data. They trained the neural network using the simulated dataset, and the input variables are module temperature, irradiance, and voltage and current at the MPP. This approach detected short-circuited modules and string disconnections. The authors tested the method experimentally and compared the performance of an ANN and a PNN. The results showed an accuracy of 90.3% for the ANN and 100% for the PNN.

Chine *et al*. [12] created a method that combines two algorithms. The first one compares the measured output power to the simulated-on MATLAB/Simulink® software. If the difference is more significant than the stipulated threshold, the algorithm identifies faults presence, and the signal enters the ANN. The RBF (Radial Basis Function) neural network was trained with a simulated dataset and diagnosis faults on bypass diodes, short-circuit and open-circuit modules, and partial shading. This method was experimentally tested by the authors and showed good accuracy.

Dhimish *et al*. [13,14] developed a fault detection for partial shading and short-circuited modules using a multilayer algorithm. The firsts layers use LabVIEW® simulations and third-order polynomial function modelling. The last layer uses a fuzzy classifier to diagnose the fault type occurring on the PV system. The method was tested using experimental data, and its results showed an accuracy of 95.27% with the fuzzy layer and 98.8% with the fuzzy layer.

Dhimish *et al*. [15] compared a fuzzy logic system to an ANN for partial shading, short-circuited module, and malfunctioning MPPT fault detection. The authors trained the RBF neural network using a voltage and power ratio, and the same variables were used to implement the Mamdani and

Sugeno fuzzy logic systems. The voltage ratio and power ratio are calculated, considering simulation results performed using MATLAB/Simulink®. The findings showed a superior accuracy of the ANN, reaching 92.1%.

Hussain *et al*. [16] compared two different ANNs for developing a fault detection method. The neural networks used were RBF and MLP (Multilayer Perceptron) for detecting disconnected PV modules on a string. The input variables were power and irradiance, and the output indicates how many faulty modules are on the string. Results showed a maximum accuracy of 97.9% on the RBF neural network.

Considering the previous discussion, it is essential to develop methods capable of identifying and diagnosing the PV system's fault. Therefore, this paper proposes a fault detection technique combining ANN and fuzzy logic to detect short-circuited modules and disconnected strings on a PV power plant. It is essential to detect this fault type since it can massively decrease power generation, and identifying it can be time-consuming, especially on large scale power plants.

A notable advantage of this work is that the proposed method is suitable and reliable once it uses pre-existing sensors, and the training dataset is obtained by simulation, not requiring long data from an existing PV system. Besides, the method does not need to compare simulated results with measured data, making it more straightforward.

The paper is briefly structured as follows. Section 2 illustrates the modelling of the PV module, explains mathematical equations needed for PV system simulation. Then, Section 3 describes the studied PV systems in this research, also validates the model simulation using experimental data. Section 4 defines the methodology used to develop the fault detection method. In Section 5, the proposed method is validated with an experimental dataset of the studied PV systems. Finally, in Section 6, the overall conclusions are discussed.

2. PV Module Modelling

Several PV cell models are proposed in the literature [17], but for this work, it was employed the one diode model, considering its simplicity. Figure 1 illustrates the equivalent circuit for the one diode model.

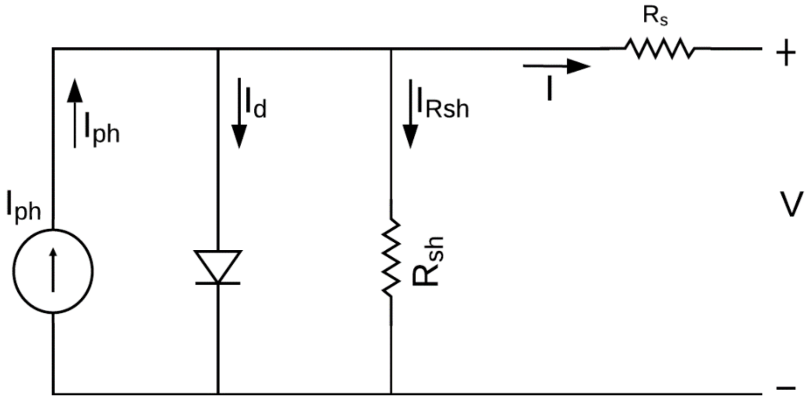


Figure 1. PV cell equivalent circuit

The circuit comprises the light-generated current (I_{ph}), parallel with a diode and a shunt resistance (R_{sh}). All these elements are series-connected to the series resistance (R_s). Analysing the circuit in Figure 1, the cell output current I can be expressed by Equation (1).

$$I = I_{ph} - I_d - I_{Rsh} \tag{1}$$

The I_d and I_{Rsh} currents represent the diode current and leakage current, respectively, and are expressed by Equations (2) and (3).

$$I_d = I_0 \left[\exp\left(\frac{V_d q}{akT}\right) - 1 \right] \tag{2}$$

$$I_{Rsh} = \frac{V + R_s I}{R_{sh}} \quad (3)$$

Substituting the I_d and I_{Rsh} expression on Equation (1), the current I delivered by the PV cell is represented on Equation (4).

$$I = I_{ph} - I_0 \left[\exp\left(\frac{V_d q}{akT}\right) - 1\right] - \frac{V + R_s I}{R_{sh}} \quad (4)$$

where:

I_0 Diode saturation current (A)
V_d Diode voltage (V)
q Electron charge ($q = 1.6 \times 10^{-19}$C)
a Diode ideality factor
k Boltzmann constant ($k = 1.38 \times 10^{-23}$J/K)
V Cell output voltage (V)
T Cell operating temperature (K)
R_s Series resistance (Ω)
R_{sh} Shunt resistance (Ω)

The light generated current I_{ph} of a PV cell depends on the irradiance and the cell operating temperature expressed by Equations (5) and (6).

$$I_{ph} = \left[I_{phn} + k_i(T - T_n)\right]\frac{G}{G_n} \quad (5)$$

$$I_{phn} = \frac{R_{sh} + R_s}{R_{sh}} I_{sc} \quad (6)$$

where:

I_{phn} Nominal light generated current (A)
I_{sc} Short-circuit current for STC (Standard Test Conditions) (A)
k_i Temperature coefficient for I_{sc} (A/K)
T_n Cell temperature for STC (298 K)
G Cell irradiance (W/m^2)
G_n Cell irradiance for STC (1000W/m^2)

The diode saturation current I_0 is related to the cell operating temperature and is expressed by Equation (7).

$$I_0 = I_{0n}\left(\frac{T_n}{T}\right)^3 \exp\left[\frac{qE_{g0}}{ak}\left(\frac{1}{T_n} - \frac{1}{T}\right)\right] \quad (7)$$

E_{g0} is the bandgap energy for semiconductor and is 1.2 eV to the polycrystalline siliceous at 25 °C [18], and the I_{0n} is the nominal saturation current, expressed by Equation (8).

$$I_{0n} = \frac{I_{sc} + k_i(T - T_n)}{\exp\left\{\frac{q[V_{oc} + k_v(T - T_n)]}{akT_n}\right\} - 1} \quad (8)$$

V_{oc} is the cell's open-circuit voltage, and k_v temperature coefficient for V_{oc} expressed in V/K. Finally, analysing the circuit in Figure 1, the diode voltage (V_d) can be represented by Equation (9).

$$V_d = V + R_s I \quad (9)$$

The one diode model characterized by Figure 1 and Equation (4) represents one single PV cell. However, in practice, a PV module comprises several connected PV cells, and a PV array comprises several connected PV modules. Thus, to analyse the I and V output characteristics of an entire PV module/array are necessary to include the parameters of the number of series-connected cells (N_s) and parallel-connected cells (N_p), as expressed by Equations (10) and (11).

$$I = N_P I_{ph} - N_P I_0 \left[\exp\left(\frac{q(V + R_s I)}{N_s akT}\right) - 1\right] - \frac{V + R_s I}{R_{sh}} \quad (10)$$

$$I_{0n} = \frac{I_{sc} + k_i(T - T_n)}{\exp\left\{\frac{q[V_{oc} + k_v(T - T_n)]}{N_s a k T_n}\right\} - 1} \quad (11)$$

It is essential to highlight that when analysing a PV module/array, R_s and R_{sh} are the equivalent resistance. Besides, V_{oc} and I_{sc} value the whole PV module/array for the Standard Test Conditions (STC). Moreover, the temperature T corresponds to the cell operating temperature, not the ambient temperature. When it is not available the cell or module temperature (T_c), it is possible to assume that T is dependent on the ambient temperature (T_a) and the Nominal Operating Cell Temperature (NOCT), as expressed by Equation (12) [19].

$$T_c = T_a + \frac{G}{800}(NOCT - 20) \quad (12)$$

Considering the model and expressions analysed, Subsection 2.1 describes PV system modelling on MATLAB/Simulink® software.

2.1. *MATLAB Simulink® Simulation*

The PV module modelling was developed using the one diode model in the MATLAB/Simulink® environment, as shown in Figure 2.

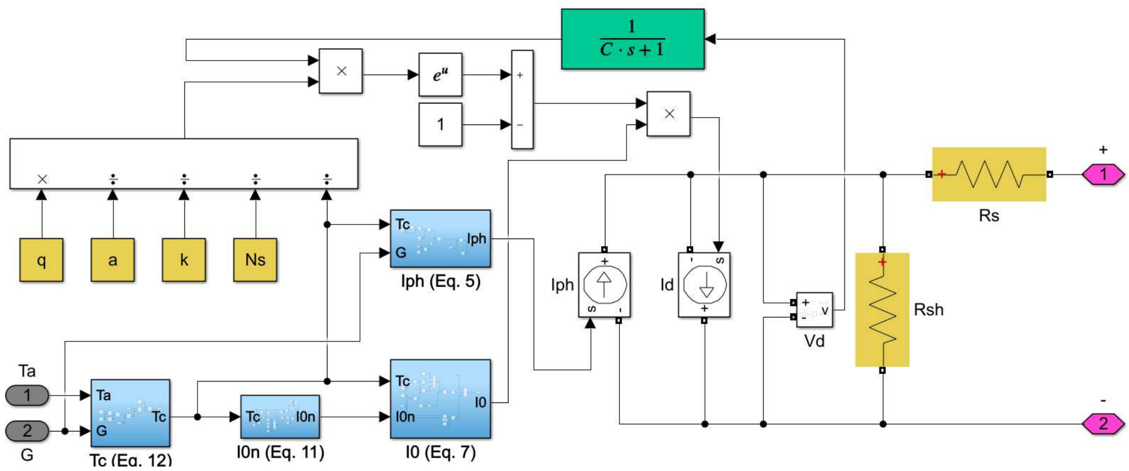


Figure 2. MATLAB/Simulink® PV module model

In Figure 2, the grey blocks are input variables, the pink blocks are the outputs of the PV modules, the yellow blocks are constants, and the blue blocks are masks containing previously discussed equations. Moreover, to avoid a loop error, it was employed a low pass filter (see the green block in Figure 2) as a feedback transfer function, and C is the filter time constant. The filter discretizes the model solution, enabling it to solve the equation and store the correct results. The time constant C should increase with the number of cells. Thus, there will be enough time for the algorithm to solve the equation, store the result, and perform the next iteration.

The manufacturers provide most of the PV modules' parameters. Generally, the parameters available on the panel datasheet are open-circuit voltage (V_{oc}), short-circuit current (I_{sc}), the Maximum Power Point (MPP) voltage (V_{MPP}), the current ate the MPP (I_{MPP}), and the power at MPP (P_{MPP}). Thus, according to Equations (10) and (11), the parameters that are not available on the PV module datasheet are the diode ideality factor (a), the series resistance (R_s), and the shunt resistance (R_{sh}). While some authors investigated how to estimate the ideality factor a [20,21], in the context of this work, it is considered $1 \le a \le 1.5$ [18]. The ideality factor a is chosen to improve the model fitting. Furthermore, the model resistances R_s and R_{sh} are calculated according to Villalva's method [18].

After modelling a PV module, it is possible to simulate an entire PV array, working under healthy or faulty conditions. The simulation enabled the development of the proposed method applied to the system described. Section 3 discusses the modelling validation.

3. Model Validation with Experimental Data

For the model validation, a comparison with experimental data is incredibly useful. It is essential to understand how the model works under different PV module models and different conditions. Thus, the proposed model will be tested for two different PV systems, named here as System 1 and System 2. Subsections 3.1 and 3.2 describes the model validation for both systems.

3.1. System 1: One String System

The PV array named System 1 in this research is illustrated in Figure 3. The system is a 2.2 kWp PV power plant, and it comprises ten series-connected PV modules. The panels model is the SMT6(60)P from PowerGlaz manufacturer, installed at the Huddersfield University campus, and Table 1 describes its characteristics.

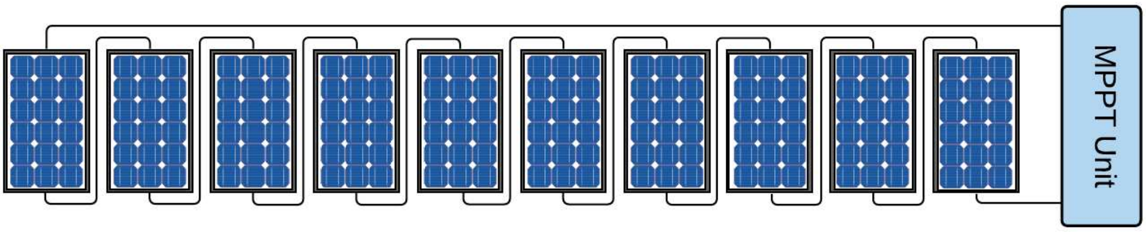
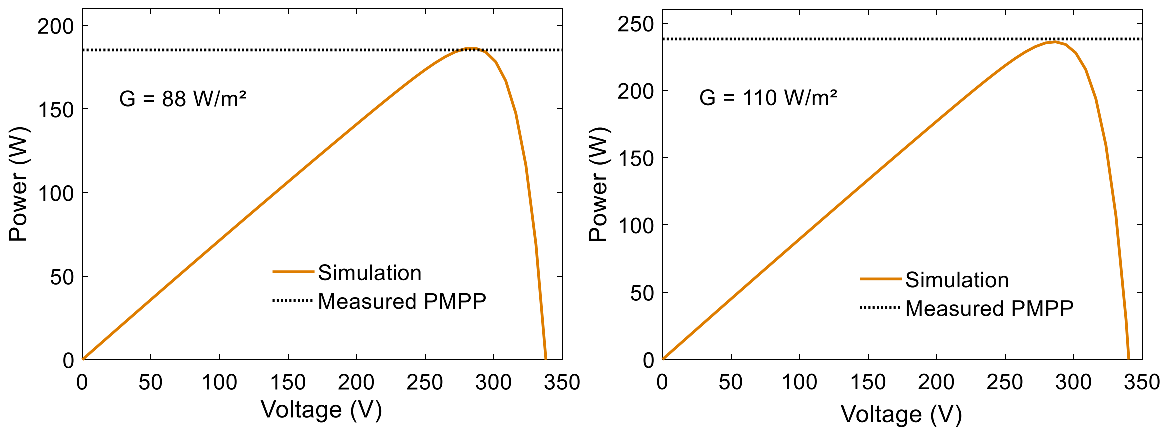


Figure 3. Schematic of System 1

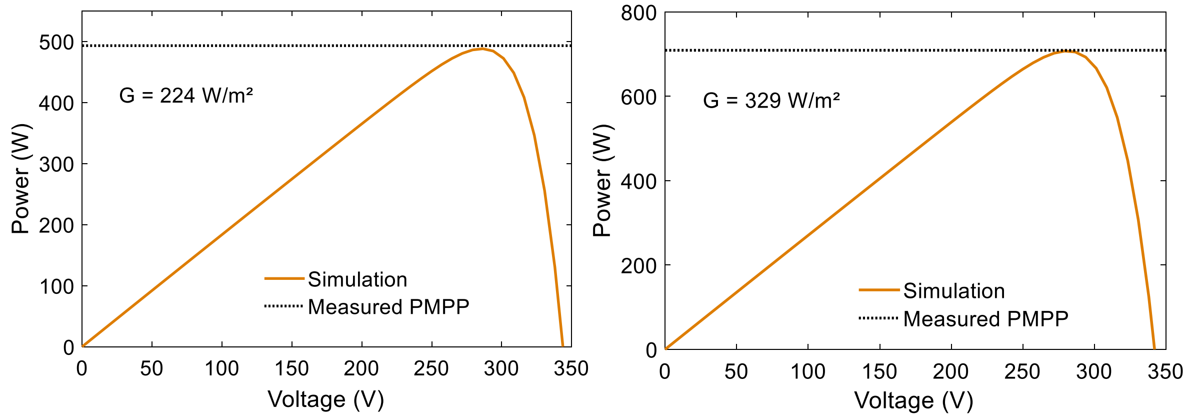
Table 1. SMT6(60)P PV modules parameters

Datasheet parameters			
Parameter	**Value**	**Parameter**	**Value**
V_{OC}	36.74 V	N_s	60
I_{SC}	8.24 A	N_p	1
k_i	0.0042 A/K	P_{MPP}	220 W
k_V	-0.132 V/K	I_{MPP}	7.7 A
NOCT	46 °C	V_{MPP}	28.7 V
Calculated parameters			
Parameter	**Value**	**Parameter**	**Value**
R_{sh} (Ω)	1108.3972	R_s (Ω)	0.3930

We simulated System 1 using the model proposed in section 2. Then, we compared the model simulation results to measured experimental data. We observed the model results varying the irradiance *G*. Figure 4 illustrates the P-V (Power vs. Voltage) curves, comparing to the experimental data.



(a) (b)



(c) (d)

Figure 4. P-V curves for System 1 (a) G = 88 W/m², (b) G = 110 W/m², (c) G = 224 W/m² and (d) G = 329 W/m²

Observing Figure 4 is possible to verify that the proposed model shows results consistent with the measured P_{MPP} for the experimented system. Table 2 summarizes a comparison of measurements of System 1 and simulation results.

Table 2. System 1 experimental results

T_a (°C)	G (W/m²)	Measured P_{MPP} (W)	Model Simulation P_{MPP} (W)	Error (%)
16	88	185.26	186.30	0.56
16	110	238.15	236.00	-0.90
16	224	493.00	487.90	-1.03
16	329	709.11	707.20	-0.27

After verifying the proposed model accuracy, we performed simulations to build the fault detection method's training database. A large dataset for the machine learning training is necessary to simulate faulty scenarios and healthy scenarios, varying the irradiance level and the module temperature. In system 1, the fault detection method is supposed to diagnose short-circuited PV modules. Thus, we simulated ten scenarios, disconnecting 1, 2, 3, until 9 modules. In each scenario, the irradiance is wide-ranging from 100 to 1100 W/m², and the ambient temperature from 10 to 40 °C. Besides, the P_{MPP} is measured for each case.

3.2. System 2: Four String System

The second PV system studied in this research, called System 2, is illustrated in Figure 5. The PV system is a 4.16 kW_P power plant and comprises 32 PV modules, arranged on four series-connected strings, with eight series-connected modules on each string. The panels model is the KC130GHT-2 from Kyocera manufacturer, also installed at the Huddersfield University campus, and Table 3 describes its characteristics.

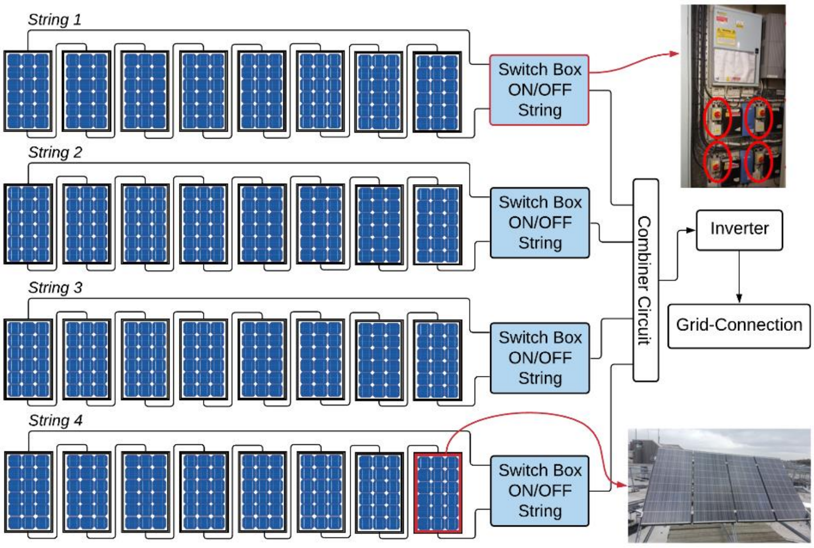
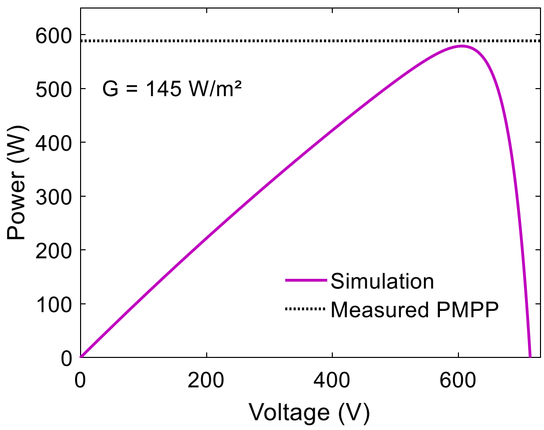


Figure 5. Schematic of System 2

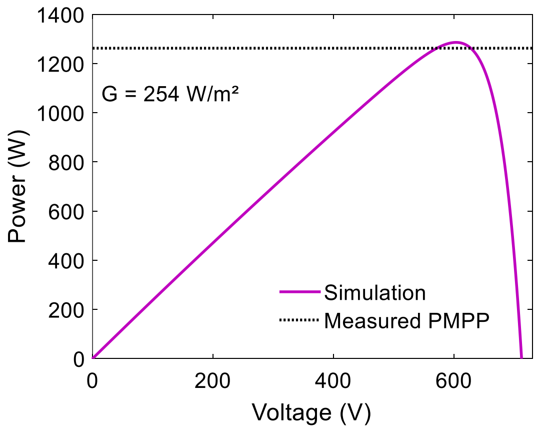
Table 3. KC130GHT-2 PV modules parameters

Datasheet parameters			
Parameter	**Value**	**Parameter**	**Value**
V_{oc}	21.90 V	N_s	36
I_{sc}	8.02 A	N_P	1
k_i	0.00318 A/K	P_{MPP}	130 W
k_v	-0.0821 V/K	I_{MPP}	7.39 A
NOCT	47 °C	V_{MPP}	17.6 V
Calculated parameters			
Parameter	**Value**	**Parameter**	**Value**
R_P (Ω)	119.232	R_s (Ω)	0.16

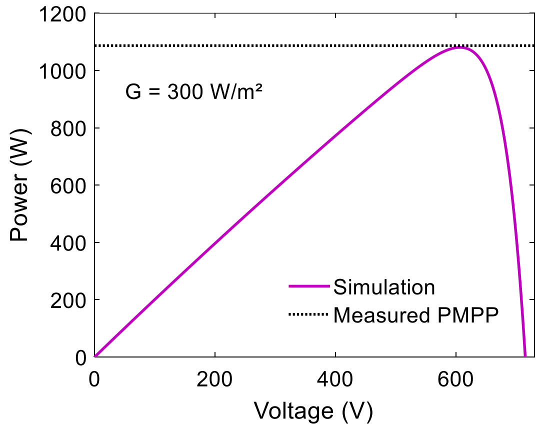
We also simulated System 2 using the model proposed in section 2. Following the same previous methodology, we compared the model simulation results to measured experimental data. We observed the model results varying the irradiance *G*. Figure 6 illustrates the P-V curves for System 2, comparing it to the experimental data.



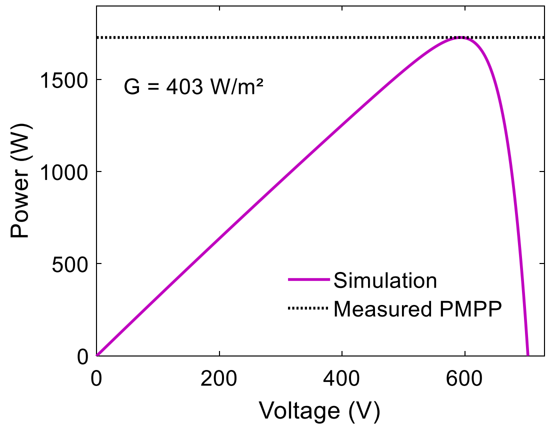
(a)



(b)



(c)



(d)

Figure 6 - P-V curves for System 1 (a) $G = 145$ W/m², (b) $G = 254$ W/m², (c) $G = 300$ W/m² and (d) $G = 403$ W/m²

Observing Figure 6 is possible to verify that the proposed model shows results consistent with the measured P_{MPP} for the experimented system. Table 4 summarizes a comparison of measurements of System 2 and simulation results.

Table 4. System 1 experimental results

T_a (°C)	G (W/m²)	Measured P_{MPP} (W)	Model Simulation P_{MPP} (W)	Error (%)
16	145	588.69	578.93	-1.66
16	254	1086.8	1080.75	0.56
16	300	1262.41	1286.00	-1.87
16	403	1701.63	1727.78	-1.54

After verifying the proposed model accuracy, we performed simulations to build the fault detection method's training database. In System 2, the fault detection method is supposed to diagnose strings disconnection fault. Thus, we modeled four scenarios, disconnecting 1, 2, and 3 strings. In each scenario, the irradiance is wide-ranging from 100 to 1100 W/m², and the ambient temperature from 10 to 40 °C. Besides, the P_{MPP} is measured for each case. With the simulated dataset, it is possible to develop the fault detection method for System 1 and 2, discussed in Section 4.

4. Fault Detection Method

The proposed fault detection method identifies short-circuited modules on System 1 and disconnected strings on System 2, indicating how many PV modules or strings are under the faulty condition. The input variables should be the irradiance (G), ambient temperature (Ta), and the measured power at the MPP (P_{MPP}). The only electrical variable, in this case, is the P_{MPP}, which makes the fault detection quite tricky. The same output power could represent various situations, including healthy and faulty conditions. Figure 7 compares two P-V curves of System 1 under different conditions to exemplify this situation.

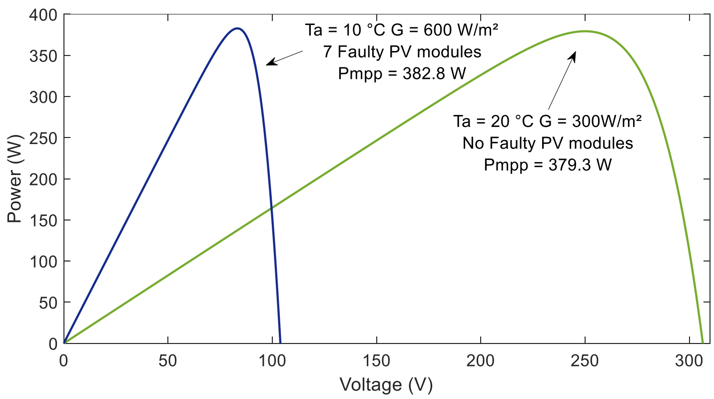


Figure 7. Comparative P-V curves of System 1 between different faulty situations

Observing Figure 7, we can see that even under entirely different conditions, the measured P_{MPP} could be quite similar. Therefore, any fault detection method needs to deal with this similarity on the database, mostly if it uses only the maximum power point (P_{MPP}) as electrical variable. Seeking to deal with this issue, we proposed combining two algorithms, as illustrated in Figure 8.

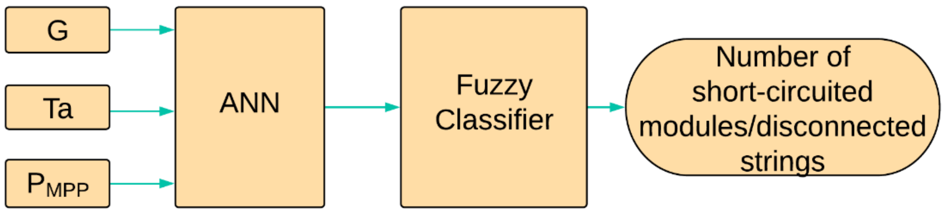


Figure 8. Fault detection method schematic

The first algorithm is an ANN using as inputs variables the irradiance G, the module temperature T_c, and the measured power at the MPP (P_{MPP}). The neural network output enters a fuzzy logic classifier that detects how many PV modules are under short-circuit fault or strings are disconnected. It is essential to highlight that the method's objective is to give the operator the exact number of short-circuited PV modules or disconnected strings on the system. Therefore, using a fuzzy classifier is essential to enable the method to deal with the similarities in the output power and still give the correct number of faults occurring on the PV system. Table 5 exemplifies the faults indicated by the detection method.

Table 5. Faults indicated by the detection method

	Short-circuited PV modules	Fault	Output
System 1	0	F0	0
	1	F1	1
	2	F2	2
	3	F3	3
	⋮	⋮	⋮
	9	F9	9
	Disconnected strings	**Fault**	**Output**
System 2	0	F0	0
	1	F1	1
	2	F2	2
	3	F3	3

System 1 comprises ten panels, so if there are ten faulty PV modules, the entire system is disconnected. Therefore, this faulty condition does not correspond to short-circuited PV modules but system failure. So, observing Table 5, the proposed method identifies 0 (normal operation) to 9 short-circuited PV modules for System 1. The ANN's and fuzzy logic details for each system are described in Subsections 4.1 and 4.2.

4.1. *Artificial Neural Network*

The ANN of the fault detection method applied to the studied system is a Multilayer Perceptron (MLP) neural network. On an MLP network, each layer has a weight matrix *W*, a bias vector *b*, and an output vector *Y*, as illustrates in Figure 9, where *f(.)* is the used activation function. The outputs of the hidden layer are defined by Equations (13) and (14) [22].

$$^{1}Y = \, ^{1}f\left(^{2}w \times X + \, ^{1}b\right) \tag{13}$$

$$^{2}Y = \, ^{2}f(^{2}w \times \, ^{1}Y + \, ^{2}b) \tag{14}$$

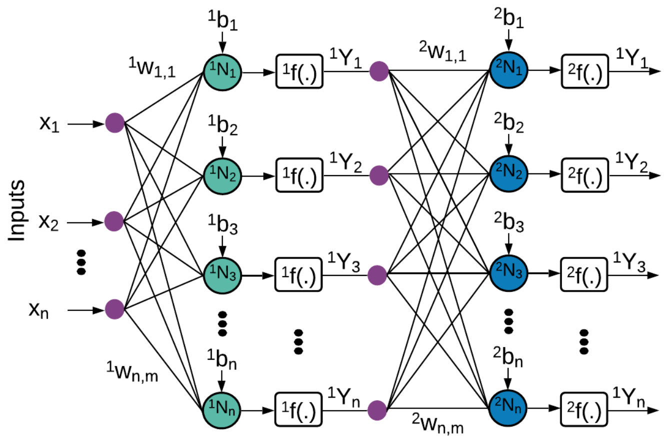


Figure 9. Basic MLP Structure

In general, MLP networks can be applied to linear or nonlinear models. Usually, it is associated with sigmoid, tansigmoid, or linear activation functions. They are often used because it provides nonzero derivatives regarding input signals and exhibits smoothness and asymptotic properties. The linear activation function is employed to approximate a continuous function in the output layer of MLP networks. There is no formal rule for choosing the number of hidden layers of neurons on it. Though, the number of neurons in the hidden layer impacts the network performance. A large number of neurons in the hidden layers will make the training process slow [22].

We developed the MLP using MATLAB® software. Figure 10 represents the structure of the MLP applied to the fault detection method, and Table 4 describes its settings.

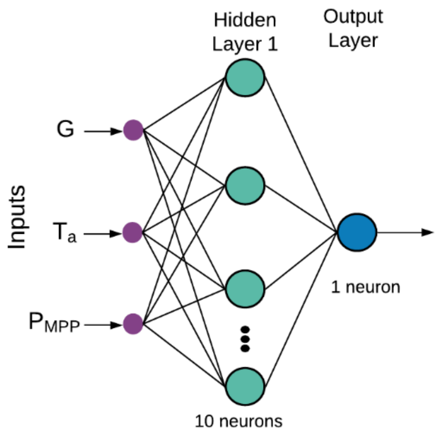


Figure 10. System 1 and 2 MLP structure

Table 6. System 1 and 2 MLP training settings

ANN Setup	
Number of input variables	3 (G, T, P_{MPP})
Number of output variables	1
Number of layers	2
Number of neurons	(10,1)
Training Process	Supervised
Training algorithm	Levenberg-Marquardt
Activation function	(tansigmoid, linear)
Training	70%
Validation	15%
Test	15%
Type of division of samples	random

The training process is supervised, meaning that we provided a set of input/output data of appropriate network behaviour. We divided randomly 70% of the samples for training, 15% for validation, and 15% for testing. Thus, we enable the validation of the desired topology. The training algorithm chosen is Levenberg-Marquardt, considering it is a faster algorithm for networks of moderate sizes.

The training dataset was obtained, as discussed in Sections 3.1 and 3.2. For System 1, it comprises 147 samples for each simulated scenario, a total of 1470 samples. For System 2, the dataset comprises 588 samples and 147 samples for each simulated scenario. We compiled the samples in a crescent order of output power (P_{MPP}), along with the respective irradiance (G), ambient temperature (T_a).

Hence, it was assumed values varying from 0 to the number of possible faults occurring on the array for the targets. Therefore, for System 1, the targets assumed ranges from 0 to 9.99, with a step 0.0068 according to the number of samples on each scenario. Thus, in training, the algorithm can understand that even for the same P_{MPP}, it could represent more than one faulty situation.

For System 2, the targets assumed ranges from 0 to 3.99. For instance, if two faulty PV modules occur on System 1, the ANN targets vary from 2 to 2.99. It is worthy of highlighting that an ANN output of 2.9 is not more critical or closer to three faulty PV modules than a 2.4 result. Both output values mean that there are two short-circuited PV modules in the system (in the case of System 1). The range on the output values is necessary to avoid incorrect fault detection in those cases of output power (P_{MPP}) are too similar even in different conditions.

Thus, each fault condition corresponds to a range of outputs values on the ANN. The training process took six epochs for both ANNs. The regression coefficients are R1 = 0.99996 and R2 = 0.99848 for System 1 and System 2 ANN's, respectively. These coefficients mean that the trained networks' outputs closely represent the ones used as training data.

The output signal is not an absolute number since each faulty condition corresponds to a range of output values, so the fuzzy logic system classifies and can determine precisely how many faults are occurring on the PV system [23].

4.2. Fuzzy Logic System

In this study, the second algorithm, combined with the ANN, is responsible for giving the operator the exact number of faulty conditions in a PV system. Considering that each faulty condition corresponds to a range of the ANN results, it could be simply trunked to the integer value by an algorithm. However, we observed that due to similarities in the P_{MPP}, as previously discussed in section 4, the ANN output not always follows the expected linearity. So, in some cases, the ANN output values are out of the range for the given faulty condition.

Therefore, considering the ANN results, a fuzzy logic system interface can precisely determine how many faulty PV modules or disconnected strings are on the examined PV system since the operator can easily set the range of the membership functions.

The implemented fuzzy logic is a Sugeno type, developed on MATLAB®, using the software's default fuzzy inference rules. . We choose the Takagi-Sugeno-Kang fuzzy inference system considering the linear relation between the inputs and outputs [24]. Figure 11 and Table 7 shows its characteristics.

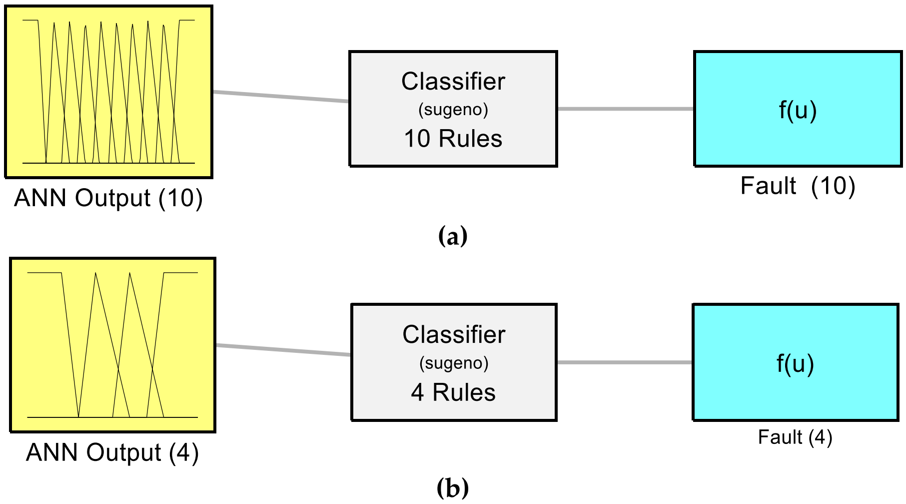


Figure 11. The developed fuzzy logic system (a) System 1 and (b) System 2

Table 7. Fuzzy logic classifier settings

Fuzzy logic system Setup	System 1	System 2
Name	Classifier	Classifier
Type	Sugeno	Sugeno
Inputs/Outputs	[1 1]	[1 1]
Number of Input Membership Functions	10	4
Number of Output Membership Functions	10	4
Number of Rules	10	4
And Method	prod	prod
Or Method	probor	probor
ImpMethod	prod	prod
AggMethod	sum	sum
DefuzzMethod	wtaver	wtaver
Input Labels	ANNoutput	ANNoutput
Output Labels	Fault	Fault
Input Range	[-1 10]	[-1 10]
Output Range	[0 1]	[0 1]
Input Membership Function Types	trimf, trapmf	trimf, trapmf
Output Membership Function Types	constant	constant

The ANN output is not an absolute number, and it enters the fuzzy classifier as an input variable. The fuzzy inference system is responsible for giving the precise number of short-circuited PV modules for System 1 and disconnected strings for System 2. Therefore, the output membership functions are constants, and Table 8 describes the input and output Membership Function (MF) settings.

Table 8. Fuzzy classifier input and output membership functions

	Input MFs		Output MFs	
	Labels	**Parameters**	**Labels**	**Parameters**
System 1	F0	[-1 -1 0 0.5]	Fault_0	[0 0 0 0]
	F1	[0.5 1 2]	Fault_1	[1 0 0 0]
	F2	[1.5 2 3]	Fault_2	[2 0 0 0]
	F3	[2.5 3 3.9]	Fault_3	[3 0 0 0]
	F4	[3.5 4 5]	Fault_4	[4 0 0 0]
	F5	[4.5 5 6]	Fault_5	[5 0 0 0]
	F6	[5.5 6 7]	Fault_6	[6 0 0 0]
	F7	[6.5 7 8]	Fault_7	[7 0 0 0]
	F8	[7.5 8 9]	Fault_8	[8 0 0 0]
	F9	[8.5 9 10 10]	Fault_9	[9 0 0 0]
System 2	F0	[-1 -1 0 0.5]	Fault_0	[0 0 0 0]
	F1	[0.5 1 2]	Fault_1	[1 0 0 0]
	F2	[1.5 2 3]	Fault_2	[2 0 0 0]
	F3	[2.5 3 3.9]	Fault_3	[3 0 0 0]

The fuzzy logic system rules are based on IF/THEN statements [25]. For the proposed fuzzy classifier, the rules are briefly listed in Table 7.

Table 9. Fuzzy classifier rules

	Fuzzy Rules
System 1	1. If (ANNoutput is F0) then (Fault is Fault_0) (1)
	2. If (ANNoutput is F1) then (Fault is Fault_1) (1)
	3. If (ANNoutput is F2) then (Fault is Fault_2) (1)
	4. If (ANNoutput is F3) then (Fault is Fault_3) (1)
	5. If (ANNoutput is F4) then (Fault is Fault_4) (1)
	6. If (ANNoutput is F5) then (Fault is Fault_5) (1)
	7. If (ANNoutput is F6) then (Fault is Fault_6) (1)
	8. If (ANNoutput is F7) then (Fault is Fault_7) (1)
	9. If (ANNoutput is F8) then (Fault is Fault_8) (1)
	10. If (ANNoutput is F9) then (Fault is Fault_9) (1)
System 2	1. If (ANNoutput is F0) then (Fault is Fault_0) (1)
	2. If (ANNoutput is F1) then (Fault is Fault_1) (1)
	3. If (ANNoutput is F2) then (Fault is Fault_2) (1)
	4. If (ANNoutput is F3) then (Fault is Fault_3) (1)

After refining the algorithms, it is attainable to test the proposed method. The following section, Section 5, discusses the testing results with experimental data.

5. Results and Discussion

In order to evaluate the effectiveness of the proposed fault detection method, the same simulated scenarios were experimentally tested. Subsections 5.1 and 5.2 describes the experimental setup and the method validation for both systems.

5.1. System 1 Experimental Setup and Method Validation

As discussed in Section 3, the PV plant comprises ten series-connected modules. The PV modules were disconnected from the string, creating all ten simulated scenarios, exemplifying the experimental setup shown in Figure 12.

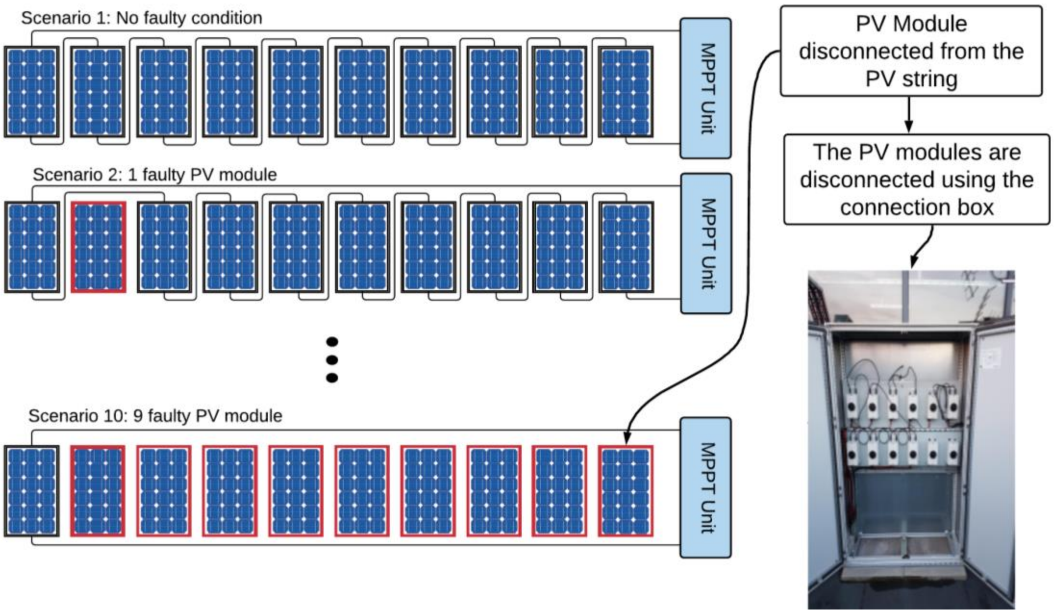


Figure 12. Experimental setup

During the experiments, the PV modules were disconnected for the entire day to collect enough data for testing the method. Although, in real situations, a faulty condition may occur not necessarily for the whole day, just for a period. The experimental tests were performed for two weeks. Figure 13 and Figure 14 depict the results. During the experiments, the irradiance (*G*), ambient temperature (T_a), and peak power (P_{MPP}) parameters were measured. The ambient temperature was constant, approximately 16 °C, in all examined days.

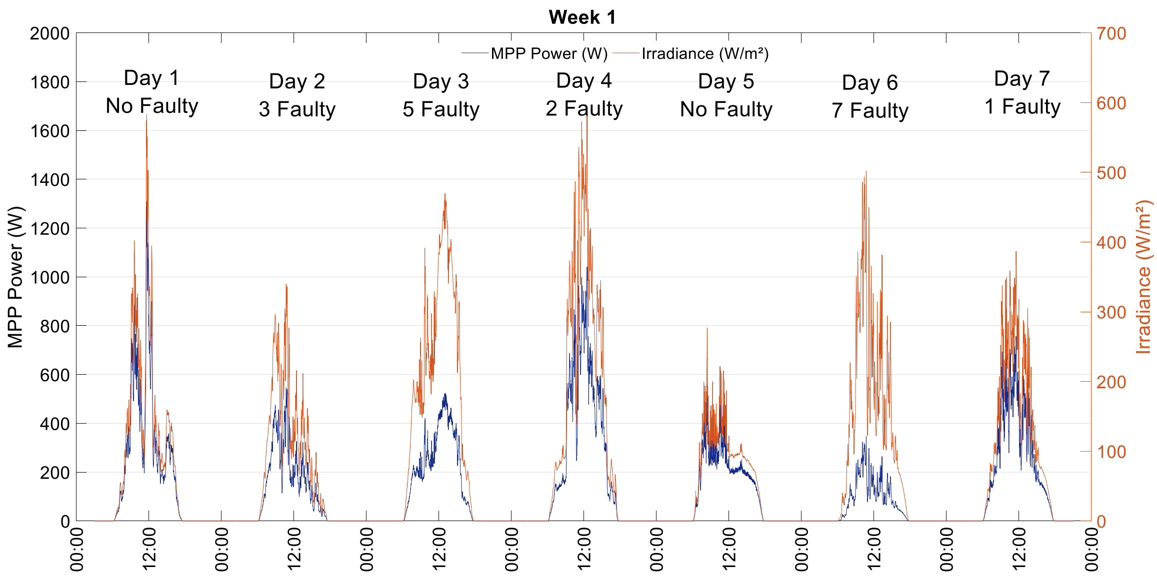


Figure 13. System 1 experimental results on week 1 for irradiance *G* and P_{MPP}

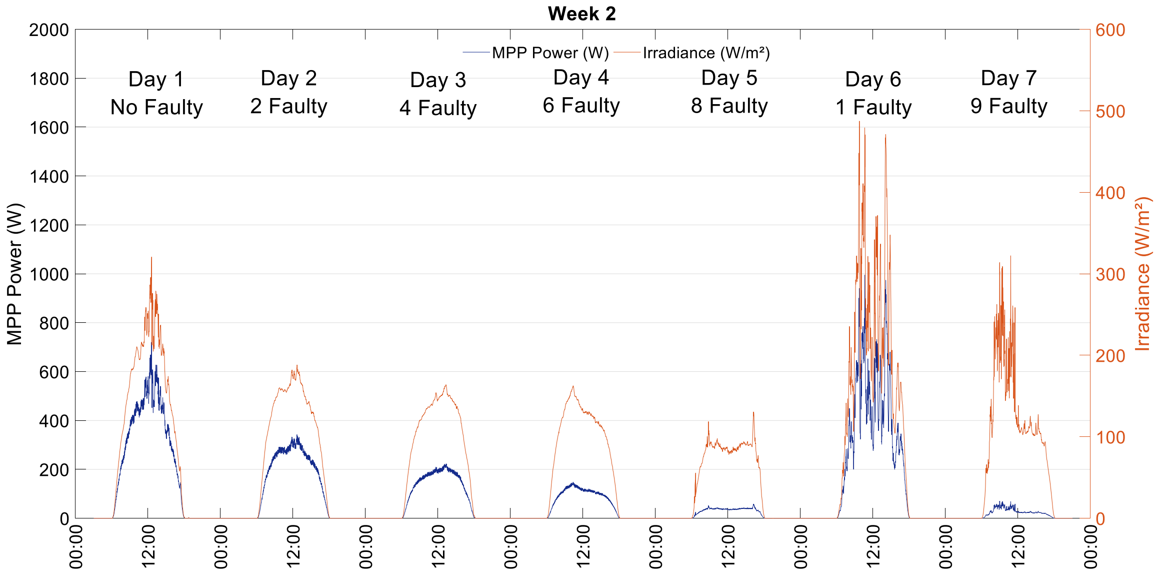


Figure 14. System 1 experimental results on week 2 for irradiance *G* and P_{MPP}

Analyzing Figure 13 and Figure 14 shows that the output power decreases significantly when a faulty situation occurs. Comparing a day with normal operation (Day 1 in Figure 13) to a faulty day (Day 7 in Figure 14), we can see that the MPP power does not follow the irradiance increase during the day, highlighting the faulty situation.

The extracted results enabled testing the proposed fault detection method. Firstly, we tested combining the ANN with a simple algorithm that truncated the ANN output to an integer value. The algorithm is responsible for giving the exact number of faulty PV modules. The truncating ranges follow the training ANN output targets (see section 4.1). Figure 15 shows the measured faulty PV modules vs. the fault detection results using the ANN combined with a truncating algorithm.

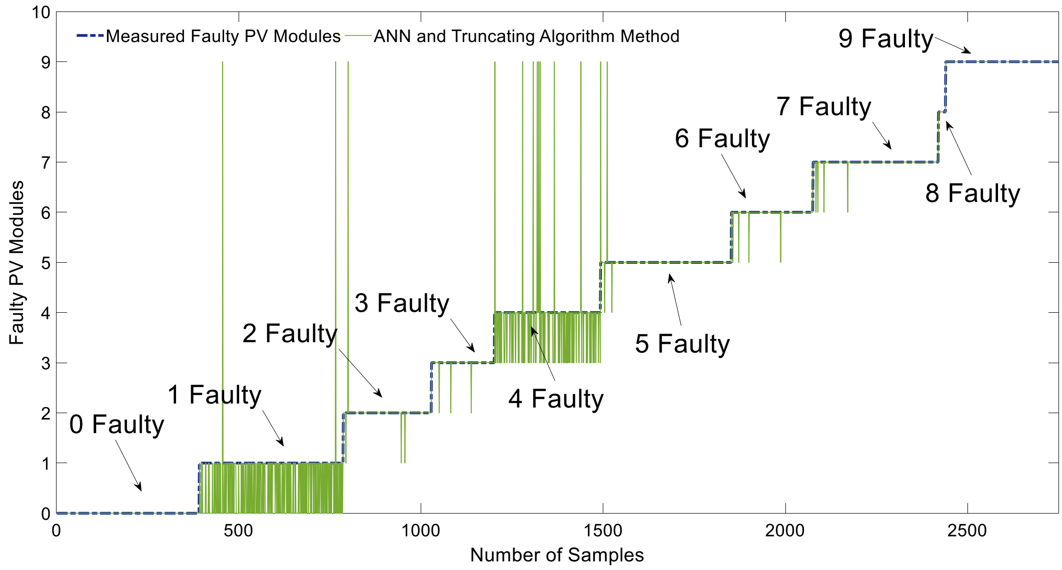


Figure 15. System 1 results using the ANN combined with a truncating algorithm

Observing Figure 15, we can conclude that combining the ANN with a simple truncating algorithm is not accurate. The critical results are on one and four faulty PV modules. Thus, combining the proposed ANN to a truncating algorithm is not suitable for fault detection.

Following, we can analyse the results of the proposed method combining the ANN and fuzzy logic system. Figure 16 shows the measured faulty PV modules vs. the neuro-fuzzy fault detection results. There is undoubtedly a significant correlation between the data points. Hence, it proves the correctness of the developed fuzzy-based system explained earlier in subsection 4.2.

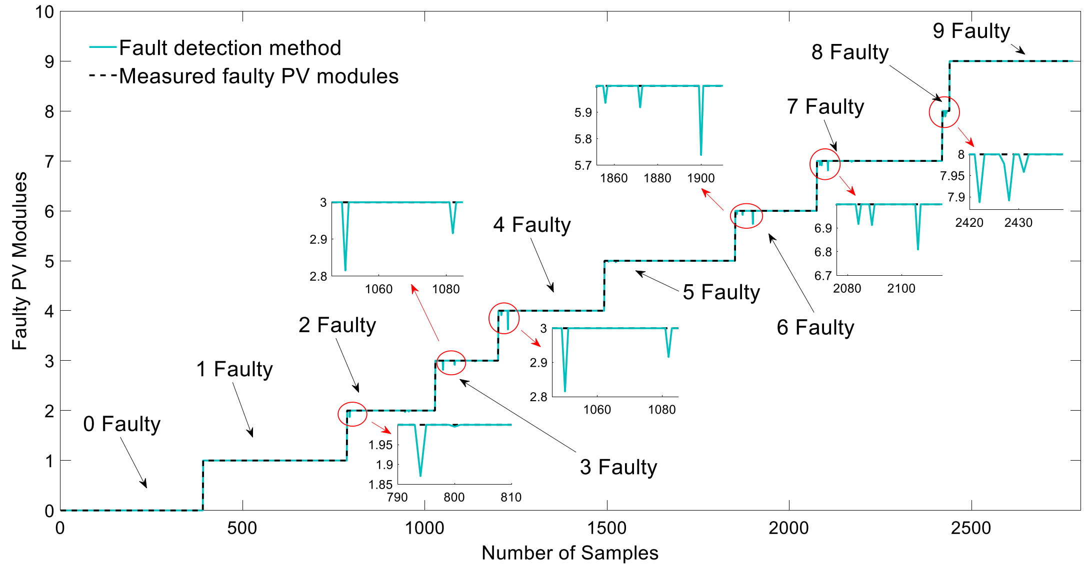


Figure 16. System 1 results using the neuro-fuzzy proposed method

The proposed method was validated using 2779 experimental samples, comprising all faulty simulated faulty situations. The lower accuracy is 98.27% for the 3 Faulty case. The weather conditions of intermittent irradiance (see Figure 13) during the experiment can justify this situation. The higher precision is observed for 0, 1, 8, and 9 Faulty cases, which achieved 100% accuracy. After all, from the results obtained, all the examined faulty conditions are accurately detected. The proposed method showed a remarkable accuracy of 99.28% for short-circuited fault detection in the studied PV system.

5.2. *System 2 Experimental Setup and Method Validation*

As discussed in Section 3.2, the PV plant comprises 32 PV modules, arranged on four strings. The strings were disconnected one at a time, using the combiner circuit box, as illustrated in Figure 5. Therefore, the experimental tests evaluated the fault case of one string disconnected. Figure 17 shows the results of 8 days of experimental tests.

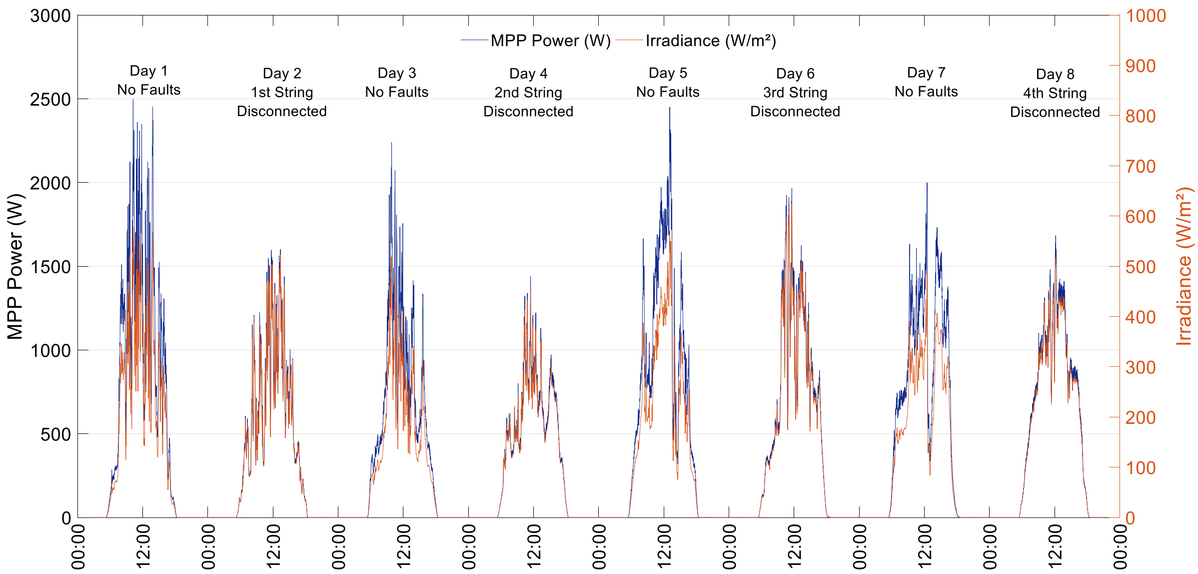


Figure 17. System 2 experimental results for irradiance *G* and P_{MPP}

During the experiments, the PV strings were disconnected for the entire day to collect enough data for testing the method, and we the irradiance (*G*), ambient temperature (T_a), and peak power (P_{MPP}). The ambient temperature was constant, approximately 16 °C, in all examined days.

Analyzing Figure 17, we can observe that the output power decreases disconnected one string. Comparing a day with normal operation (Day 1) to a one disconnected string (Day 5), we observe that the MPP power does not follow the irradiance increase during the day, highlighting the faulty situation.

The extracted results enabled testing the proposed fault detection method. For System 2, we also tested combining the ANN with a truncating algorithm. In this case, the algorithm is responsible for round the ANN output and gives the exact number of disconnected strings on the system. Figure 18 shows the measured faulty PV modules vs. the fault detection results using the ANN combined with a truncating algorithm.

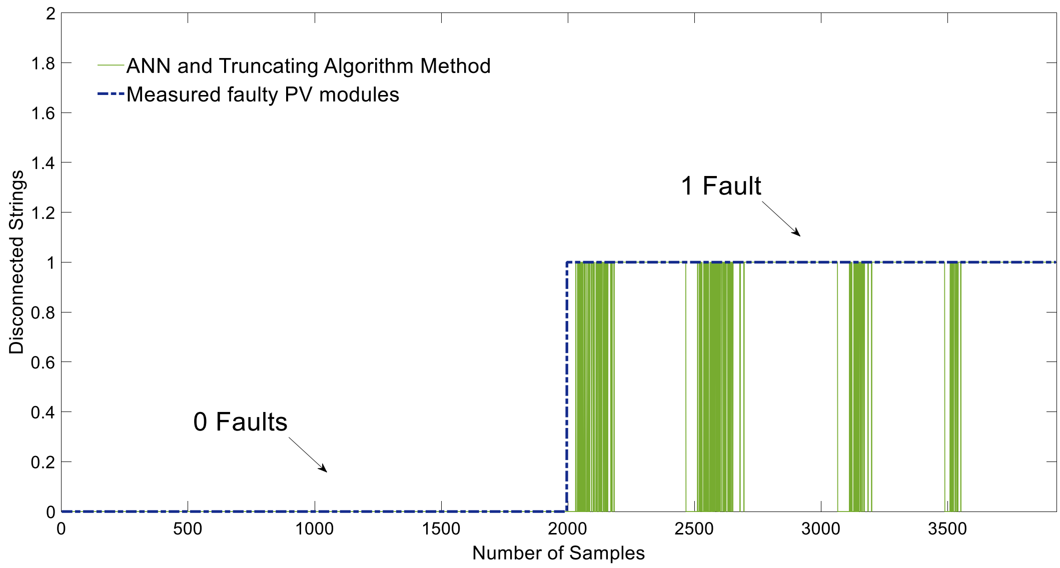


Figure 18. System 2 results using the ANN combined with a truncating algorithm

Analysing Figure 18, we can conclude that combining the ANN with a simple truncating algorithm is not accurate for System 2, just like happened to System 1. Thus, combining the proposed ANN to a truncating algorithm is not suitable for fault detection.

Following, we can analyse the results of the proposed method combining the ANN and fuzzy logic system. Figure 19 shows the measured faulty PV modules vs. the neuro-fuzzy fault detection results for System 2. Following System 1 results, there is undoubtedly a significant correlation between the data points. Hence, the accuracy of the developed fuzzy-based system explained earlier in subsection 4.2.

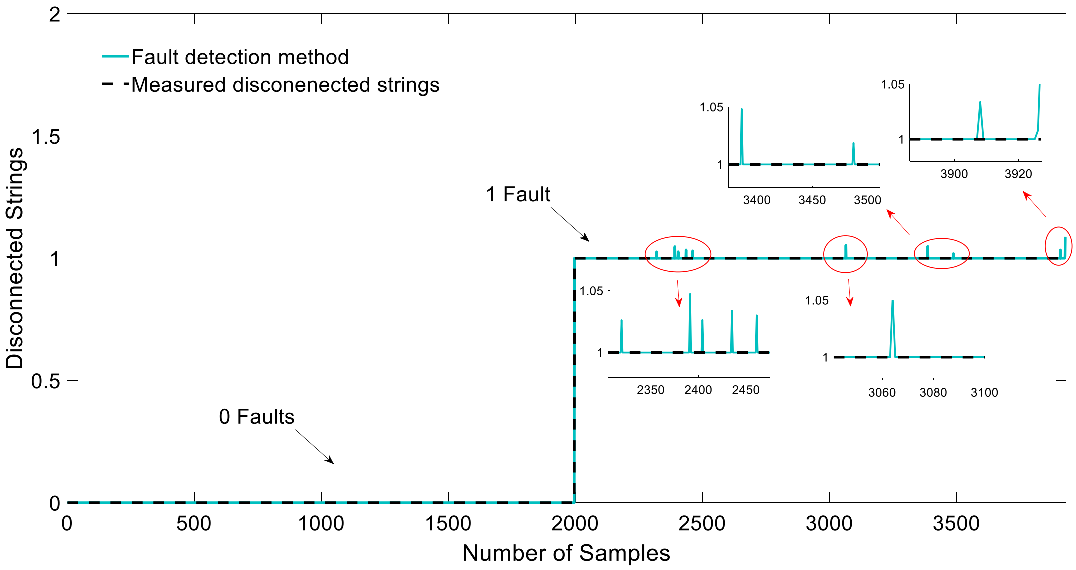


Figure 19. System 2 results using the neuro-fuzzy proposed method

For System 2, the proposed method was validated using 3927 measured samples, it comprises regular operation and one string disconnected. The tests with the experimental dataset showed an accuracy of 99.43% identifying string disconnection. These findings allow us to conclude that the proposed method is remarkably useful in detecting fault conditions on PV systems. After validating the proposed model, Section 6 discussed the overall conclusion of this research.

6. Conclusions

This paper proposes a reliable and straightforward method for fault detection on PV systems, detecting short-circuited PV modules, and string disconnection. The method comprises two machine learning algorithms. The first one is an ANN, and the second a fuzzy logic inference system. The ANN is a multilayer feedforward neural network, and the training process used a simulated dataset. Therefore, it makes the method applicable to any PV plant, also does not require long datasets from pre-existing systems. The input variables are irradiance, ambient temperature, and power at the maximum power point. The ANN output enters a Sugeno type fuzzy logic classifier, precisely determining how many short-circuited PV modules are on the given PV array.

The proposed method was validated using experimental data from two different PV systems installed on the Huddersfield University campus. The first one, named here as System 1, comprises a 2.2 kWp PV system. The obtained results for System 1 showed a remarkable accuracy of 99.28%. The second system, named System 2, is a 4.16 kW_P PV system. The obtained results, in this case, showed an accuracy of 99.43%.

These findings allowed us to conclude that the proposed method, combing ANN, and fuzzy logic systems, is accurate for detecting short-circuited PV modules and disconnected strings. Besides, it is worthy of highlighting that the proposed method does not require installing any different sensors than those that already exist on a large PV power plant, and it possible to apply to any PV system. Thus, it makes it easier for implementing the proposed method.

Author Contributions: R. G. V. conceived the methodology, developed theory, performed simulations, and validated the method. M.D. provided experimental results for model validation and helped to shape the research. F. M. U. A. conceived the idea, performed the supervision, and contributed to the revision of the manuscript. M. I. S. G. provided critical feedback and contributed to the revision of the manuscript.

Funding: This research received no external funding

Acknowledgments: The authors would like to acknowledge the support provided by the Semi-Arid Federal University, the University of Rio Grande do Norte, and the University of Huddersfield in the framework of an international contribution.

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

References

1. IEA *Snapshot of Global Photovoltaic Markets*; Brussels, 2020;
2. Ghaffarzadeh, N.; Azadian, A. A Comprehensive Review and Performance Evaluation in Solar (PV) Systems Fault Classification and Fault Detection Techniques. *J. Sol. Energy Res.* **2019**, *4*, 252–272.
3. Pillai, D.S.; Blaabjerg, F.; Rajasekar, N. A Comparative Evaluation of Advanced Fault Detection Approaches for PV Systems. *IEEE J. Photovoltaics* **2019**, *9*, 513–527.
4. Madeti, S.R.; Singh, S.N. A Comprehensive Study on Different Types of Faults and Detection Techniques for Solar Photovoltaic System. *Sol. Energy* **2017**, *158*, 161–185.
5. Dhimish, M.; Chen, Z. Novel Open-Circuit Photovoltaic Bypass Diode Fault Detection Algorithm. IEEE J. Photovolt. 2019, 9, 1819–1827.
6. Leva, S.; Mussetta, M.; Ogliari, E. PV Module Fault Diagnosis Based on Microconverters and Day-Ahead Forecast. *IEEE Trans. Ind. Electron.* **2019**, *66*, 3928–3937.
7. 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; IEEE, 25-28 September: Hersonissos, Greece, 2011; pp. 1–6.
8. Li, Z.; Wang, Y.; Zhou, D.; Wu, C. An Intelligent Method for Fault Diagnosis in Photovoltaic Array. In Proceedings of the International Conference on Electrical and Information Technologies (ICEIT); IEEE, 15-18 November: Rabat, Morocco, 2017; pp. 10–16.
9. Jiang, L.L.; Maskell, D.L. Automatic Fault Detection and Diagnosis for Photovoltaic Systems Using Combined Artificial Neural Network and Analytical Based Methods. In Proceedings of the Proceedings of the International Joint Conference on Neural Networks; IEEE, 12-17 July: Killarney, Ireland, 2015; Vol. 2015-Septe.
10. Akram, M.N.; Lotfifard, S. Modeling and Health Monitoring of DC Side of Photovoltaic Array. *IEEE Trans. Sustain. Energy* **2015**, 1–9.
11. 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.
12. Chine, W.; Mellit, A.; Lughi, V.; Malek, A.; Sulligoi, G.; Massi Pavan, A. A Novel Fault Diagnosis Technique for Photovoltaic Systems Based on Artificial Neural Networks. *Renew. Energy* **2016**, *90*, 501–512.
13. 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.
14. Dhimish, M.; Holmes, V.; Mehrdadi, B.; Dales, M. Multi-Layer Photovoltaic Fault Detection Algorithm. *High Volt.* **2017**, *2*, 244–252.
15. Dhimish, M.; Holmes, V.; Mehrdadi, B.; Dales, M. Comparing Mamdani Sugeno fuzzy logic and RBF ANN network for PV fault detection. *Renew. Energy* **2018**, *117*, 257–274.
16. 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.
17. Benghanem, M.S.; Alamri, S.N. Modeling of Photovoltaic Module and Experimental Determination of Serial Resistance. *J. Taibah Univ. Sci.* **2009**, *2*, 94–105.
18. Villalva, M.G.; Gazoli, J.R.; Filho, and E.R. Comprehensive Approach to Modeling and Simulation of

Photovoltaic Arrays. *IEEE Trans. POWER Electron.* **2009**, *24*, 1198–1208.

19. Khatib, T.; Elmenreich, W. *Modeling of Photovoltaic Systems Using MATLAB*; 1st ed.; John Wiley & Sons: Hoboken, 2016; Vol. 3; ISBN 9781119118107.
20. Bayhan, H.; Bayhan, M. A Simple Approach to Determine the Solar Cell Diode Ideality Factor Under Illumination. *Sol. Energy* **2011**, *85*, 769–775.
21. Meyer, E.L. Extraction of Saturation Current and Ideality Factor from Measuring Voc and Isc of Photovoltaic Modules. *Int. J. Photoenergy* **2017**, *2017*.
22. Siddique, N.; Adeli, H. *Computational Intelligence Computational Intelligence Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing*; 1st ed.; John Wiley & Sons: New Delhi, India, 2013; ISBN 9781118337844.
23. Dhimish, M.; Badran, G. Photovoltaic Hot-Spots Fault Detection Algorithm Using Fuzzy Systems. *IEEE Trans. Device Mater. Reliab.* **2019**, *19*, 671–679.
24. Engelbrecht, A.P. *Computational intelligence: An introduction*; 2nd ed.; Jonh Wiley & Sons, Ltd: Chichester, West Sussex, United Kingdom, 2007; Vol. 1;
25. Dhimish, B.M.; Holmes, V.; Mehrdadi, B.; Dales, M.; Mather, P. Detecting Defective Bypass Diodes in Photovoltaic Modules using Mamdani Fuzzy Logic System. Glob. J. Res. Eng. Electr. Electron. Eng. 2017, 17, 33–42.