



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/177725/>

Version: Accepted Version

Article:

Vieira, Romênia G., Dhimish, Mahmoud, Araújo, Fábio M. U. de et al. (2020) PV Module Fault Detection Using Combined Artificial Neural Network and Sugeno Fuzzy Logic. Electronics. 2150. ISSN: 2079-9292

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

1 Article

2 PV Module Fault Detection Using Combined 3 Artificial Neural Network and Sugeno Fuzzy Logic

4 Romênia G. Vieira ¹, Mahmoud Dhimish ^{2*}, Fábio M. U. de Araújo ³ and Maria I. S. Guerra ¹

5 ¹ Department of Engineering and Technology, Semi-Arid Federal University, Francisco Mota Av., Mossoro,
6 59625-900, Brazil; romenia.vieira@ufersa.edu.br (R.G.V); izabel.guerra@ufersa.edu.br (M.I.S.G)

7 ² Department of Engineering and Technology, University of Huddersfield, Huddersfield, HD1 3DH, United
8 Kingdom

9 ³ Department of Computer and Automation Engineering, Federal University of Rio Grande do Norte, Natal,
10 59078-970, Brazil; meneghet@dca.ufrn.br

11 * Correspondence: M.A.Dhimish@hud.ac.uk

12 Received: date; Accepted: date; Published: date

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

22 **Keywords:** Photovoltaic system; Photovoltaic faults; Fault detection; ANN networks; Fuzzy logic
23 system.
24

25 1. Introduction

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

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

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

40 There are various PV modules faults sources, such as mismatch, bypass diode, circuit faults,
41 asymmetrical faults, arc faults, ground faults, and lightning. It can be temporary or permanent,
42 depending on the cause and the period that affects the PV systems performance [4]. The circuit faults,
43 which are the subject of this research, can be open-circuit or short-circuit. In both situations, the PV

44 module does not contribute to power generation, so it could significantly decrease the system
45 performance and remain undetected.

46 Therefore, some monitoring and diagnosis techniques have been developed in recent years.
47 Diagnosis methods can be classified into two general groups: visual and electric. Visual methods
48 require constant verification by an operator and some specific types of equipment. On the other hand,
49 electrical methods use output signals such as voltage, current, and power to identify faults occurrence
50 [5]. Besides, it is possible to use existing sensors and pieces of equipment, making it more viable for
51 implementing the fault detection method.

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

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

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

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

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

88 Dhimish *et al.* [13,14] developed a fault detection for partial shading and short-circuited modules
89 using a multilayer algorithm. The firsts layers use LabVIEW® simulations and third-order polynomial
90 function modelling. The last layer uses a fuzzy classifier to diagnose the fault type occurring on the
91 PV system. The method was tested using experimental data, and its results showed an accuracy of
92 95.27% with the fuzzy layer and 98.8% with the fuzzy layer.

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

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

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

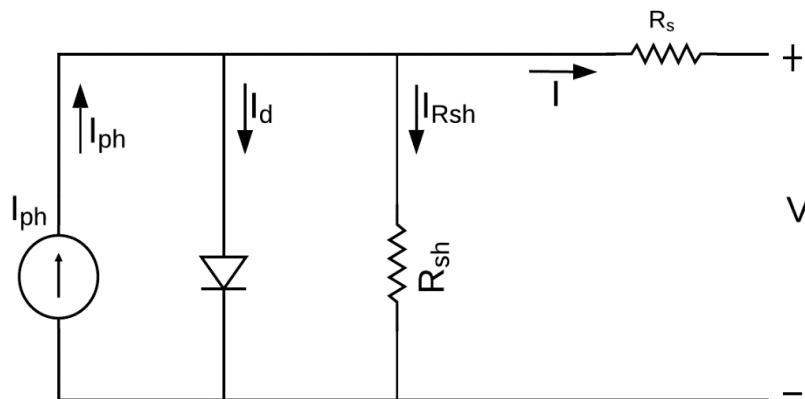
104 Considering the previous discussion, it is essential to develop methods capable of identifying
 105 and diagnosing the PV system's fault. Therefore, this paper proposes a fault detection technique
 106 combining ANN and fuzzy logic to detect short-circuited modules and disconnected strings on a PV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

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

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

119 2. PV Module Modelling

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



123

124

Figure 1. PV cell equivalent circuit

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

$$I = I_{ph} - I_d - I_{Rsh} \quad (1)$$

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

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

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

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

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

132 where:

133 I_0 Diode saturation current (A)

134 V_d Diode voltage (V)

135 q Electron charge ($q = 1.6 \times 10^{-19}$ C)

136 a Diode ideality factor

137 k Boltzmann constant ($k = 1.38 \times 10^{-23}$ J/K)

138 V Cell output voltage (V)

139 T Cell operating temperature (K)

140 R_s Series resistance (Ω)

141 R_{sh} Shunt resistance (Ω)

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

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

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

144 where:

145 I_{phn} Nominal light generated current (A)

146 I_{sc} Short-circuit current for STC (Standard Test Conditions) (A)

147 k_i Temperature coefficient for I_{sc} (A/K)

148 T_n Cell temperature for STC (298 K)

149 G Cell irradiance (W/m²)

150 G_n Cell irradiance for STC (1000W/m²)

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

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

153 E_{g0} is the bandgap energy for semiconductor and is 1.2 eV to the polycrystalline siliceous at 25
154 °C [18], and the I_{0n} is the nominal saturation current, expressed by Equation (8).

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

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

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

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

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

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

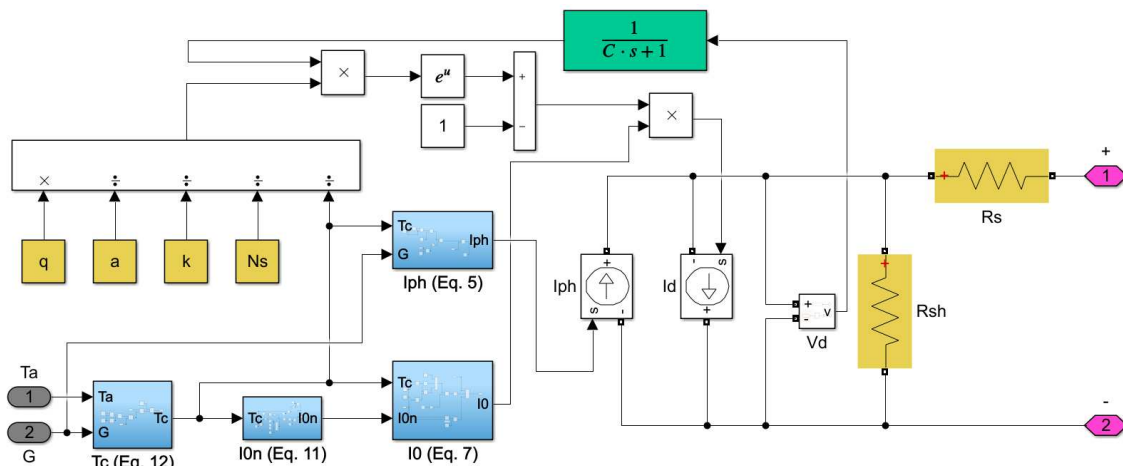
162 It is essential to highlight that when analysing a PV module/array, R_s and R_{sh} are the equivalent
 163 resistance. Besides, V_{oc} and I_{sc} value the whole PV module/array for the Standard Test Conditions
 164 (STC). Moreover, the temperature T corresponds to the cell operating temperature,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

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

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

170 2.1. MATLAB Simulink® Simulation

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



173
174 **Figure 2.** MATLAB/Simulink® PV module model

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

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

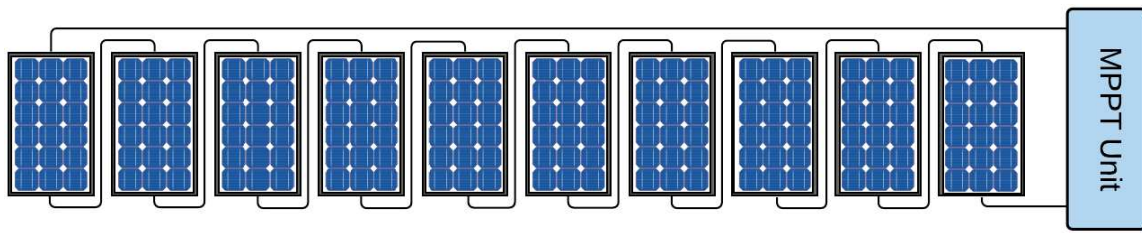
190 After modelling a PV module, it is possible to simulate an entire PV array, working under
 191 healthy or faulty conditions. The simulation enabled the development of the proposed method
 192 applied to the system described. Section 3 discusses the modelling validation.

193 **3. Model Validation with Experimental Data**

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

198 **3.1. System 1: One String System**

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

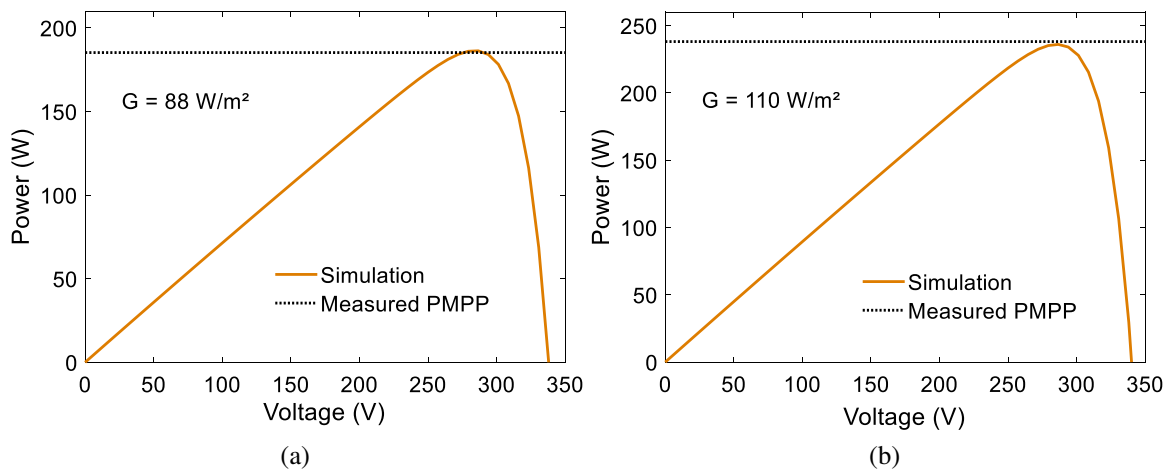


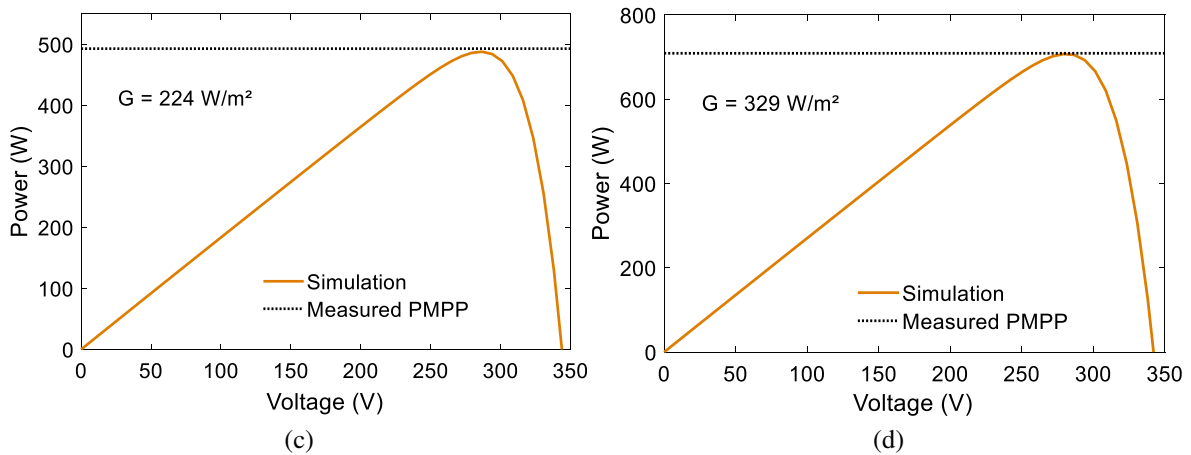
203
204 **Figure 3.** Schematic of System 1

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
$R_{sh} (\Omega)$	1108.3972	$R_s (\Omega)$	0.3930

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





209

Figure 4. P-V curves for System 1 (a) $G = 88 \text{ W/m}^2$, (b) $G = 110 \text{ W/m}^2$, (c) $G = 224 \text{ W/m}^2$ and (d) $G = 329 \text{ W/m}^2$

210

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.

211

212

Table 2. System 1 experimental results

T_a (°C)	G (W/m ²)	Measured	Model Simulation	Error (%)
		P_{MPP} (W)	P_{MPP} (W)	
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

213

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.

214

215

216

217

218

219

220

3.2. System 2: Four String System

221

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.

222

223

224

225

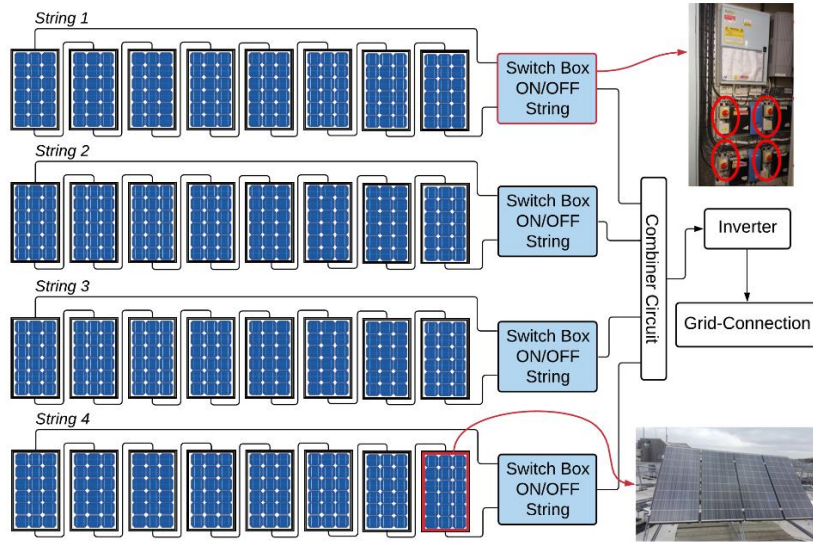
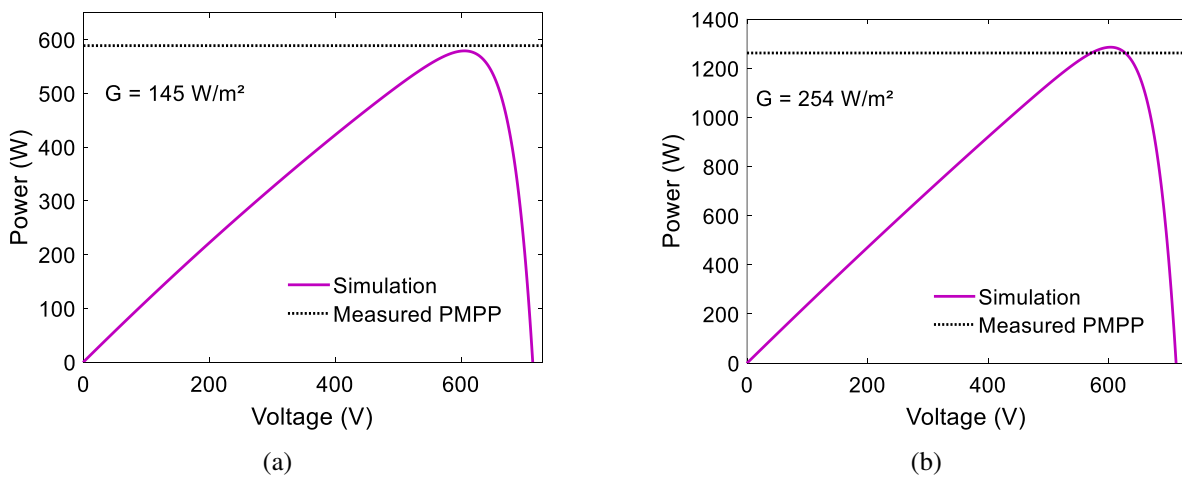


Figure 5. Schematic of System 2

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

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



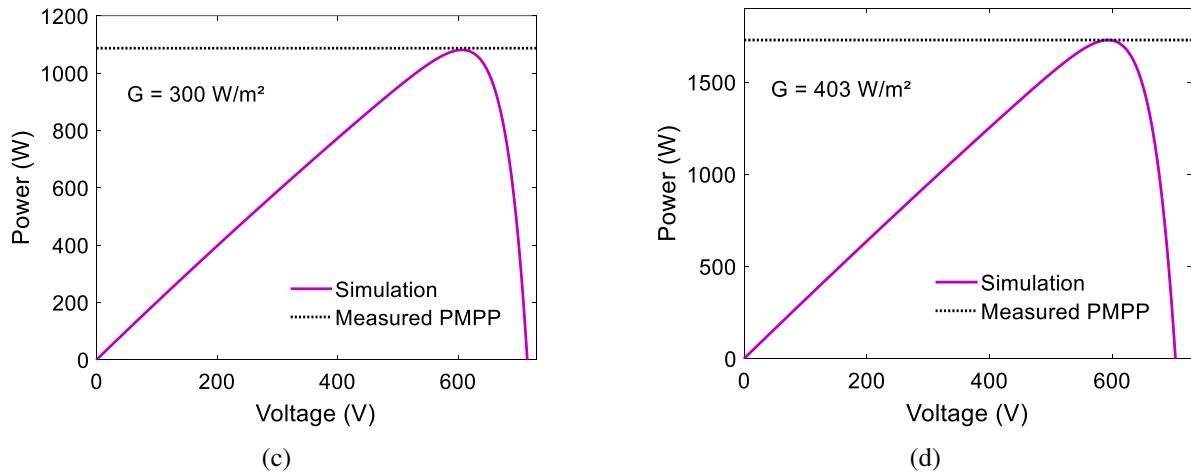


Figure 6 - P-V curves for System 1 (a) $G = 145 \text{ W/m}^2$, (b) $G = 254 \text{ W/m}^2$, (c) $G = 300 \text{ W/m}^2$ and (d) $G = 403 \text{ W/m}^2$

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

Table 4. System 1 experimental results

T_a (°C)	G (W/m ²)	Measured	Model Simulation	Error (%)
		P_{MPP} (W)	P_{MPP} (W)	
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

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

239 4. Fault Detection Method

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

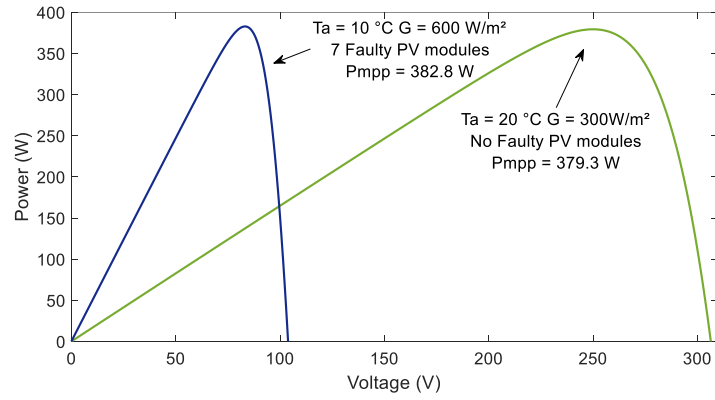


Figure 7. Comparative P-V curves of System 1 between different faulty situations

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

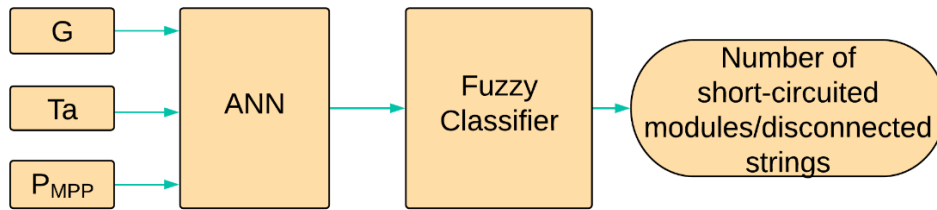


Figure 8. Fault detection method schematic

251 The first algorithm is an ANN using as inputs variables the irradiance G , the module
 252 temperature T_c , and the measured power at the MPP (P_{MPP}). The neural network output enters a fuzzy
 253 logic classifier that detects how many PV modules are under short-circuit fault or strings are
 254 disconnected. It is essential to highlight that the method's objective is to give the operator the exact
 255 number of short-circuited PV modules or disconnected strings on the system. Therefore, using a
 256 fuzzy classifier is essential to enable the method to deal with the similarities in the output power and
 257 still give the correct number of faults occurring on the PV system. Table 5 exemplifies the faults
 258 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

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

264 4.1. Artificial Neural Network

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

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

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

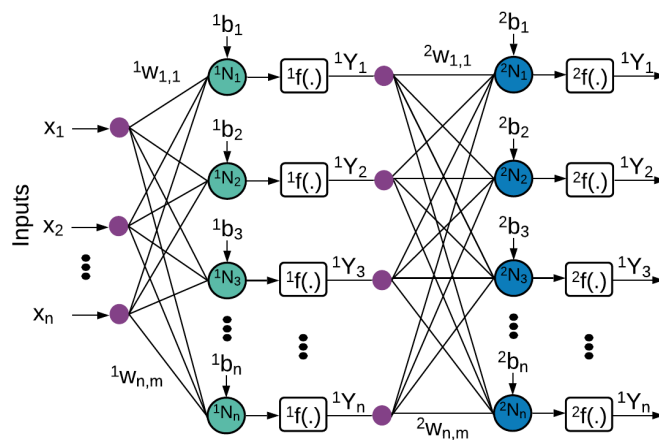


Figure 9. Basic MLP Structure

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

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

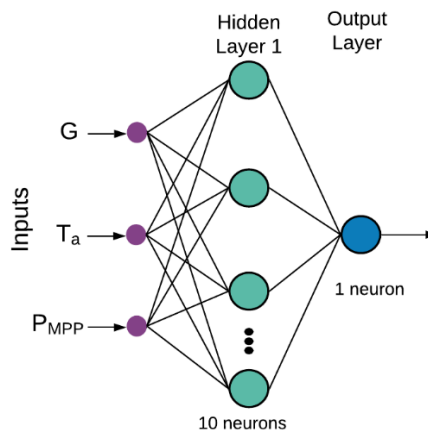


Figure 10. System 1 and 2 MLP structure

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

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

283 The training dataset was obtained, as discussed in Sections 3.1 and 3.2. For System 1, it comprises
 284 147 samples for each simulated scenario, a total of 1470 samples. For System 2, the dataset comprises
 285 588 samples and 147 samples for each simulated scenario. We compiled the samples in a crescent
 286 order of output power (P_{MPP}), along with the respective irradiance (G), ambient temperature (T_a).

287 Hence, it was assumed values varying from 0 to the number of possible faults occurring on the
 288 array for the targets. Therefore, for System 1, the targets assumed ranges from 0 to 9.99, with a step
 289 0.0068 according to the number of samples on each scenario. Thus, in training, the algorithm can
 290 understand that even for the same P_{MPP} , it could represent more than one faulty situation.

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

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

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

304 4.2. Fuzzy Logic System

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

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

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

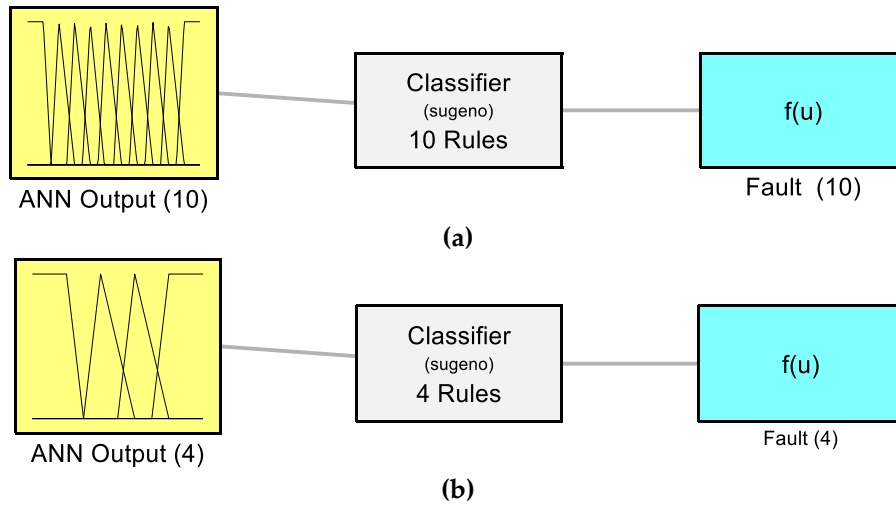


Figure 11. The developed fuzzy logic system (a) System 1 and (b) System 2

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

318 The ANN output is not an absolute number, and it enters the fuzzy classifier as an input variable.
 319 The fuzzy inference system is responsible for giving the precise number of short-circuited PV
 320 modules for System 1 and disconnected strings for System 2. Therefore, the output membership
 321 functions are constants, and Table 8 describes the input and output Membership Function (MF)
 322 settings.
 323

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]

324 The fuzzy logic system rules are based on IF/THEN statements [25]. For the proposed fuzzy
 325 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)

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

328 **5. Results and Discussion**

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

332 *5.1. System 1 Experimental Setup and Method Validation*

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

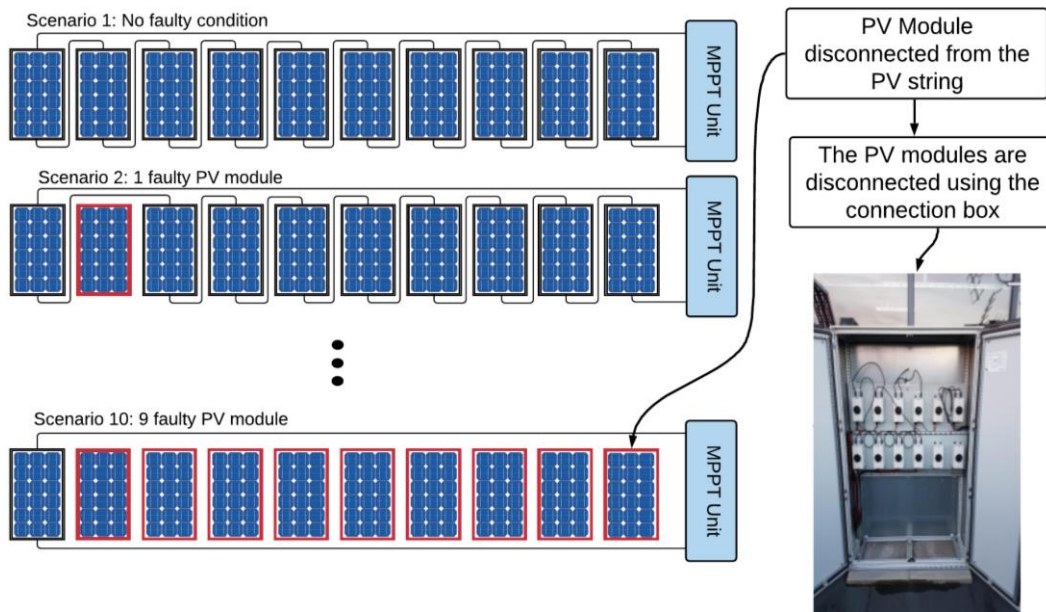


Figure 12. Experimental setup

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

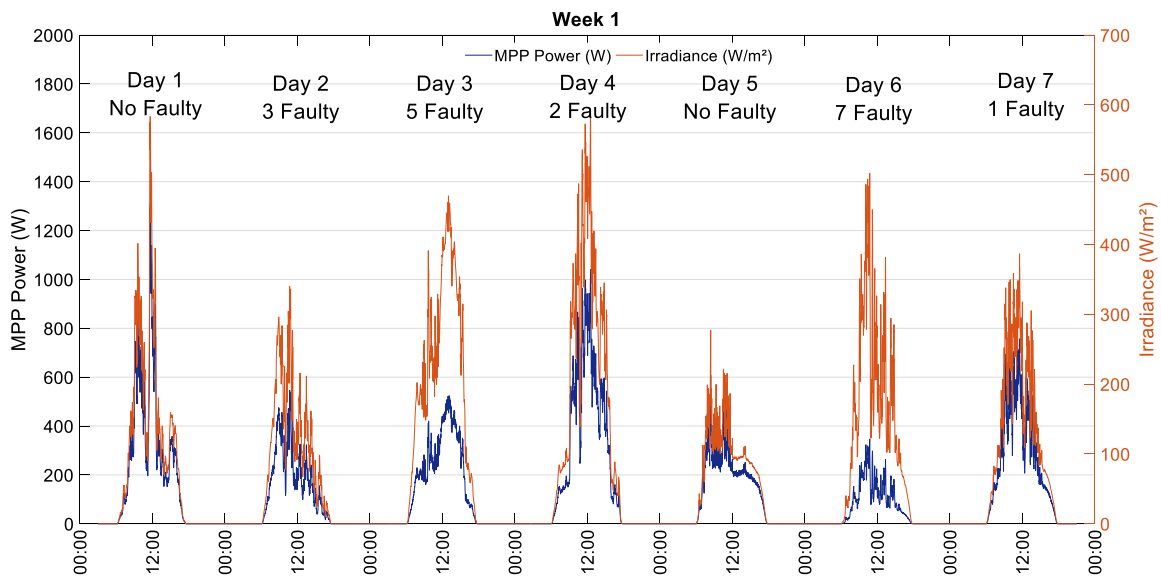


Figure 13. System 1 experimental results on week 1 for irradiance G and P_{MPP}

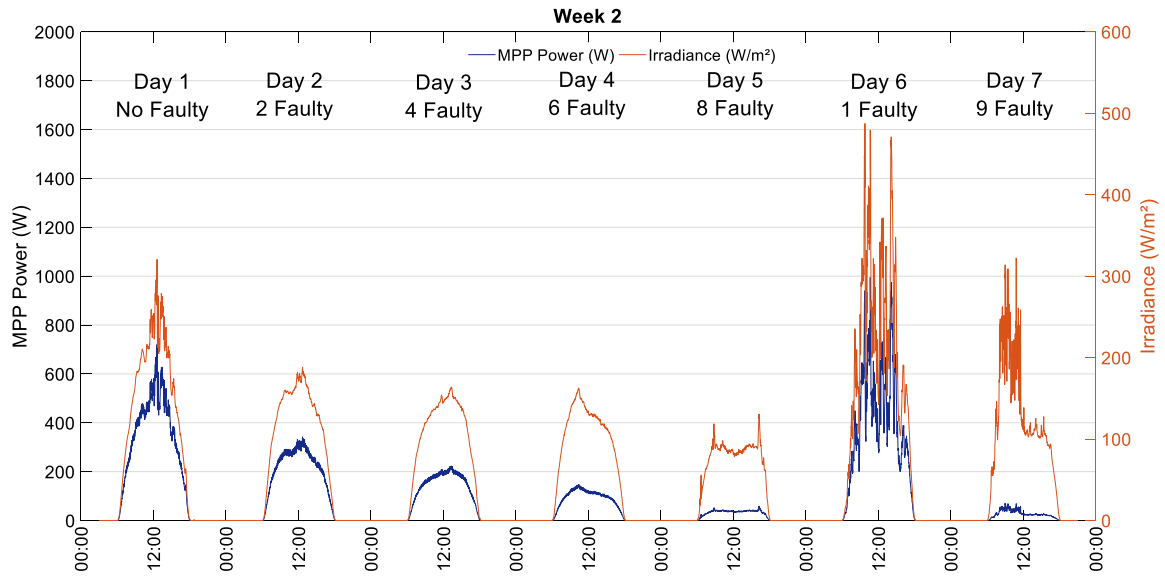


Figure 14. System 1 experimental results on week 2 for irradiance G and P_{MPP}

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

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

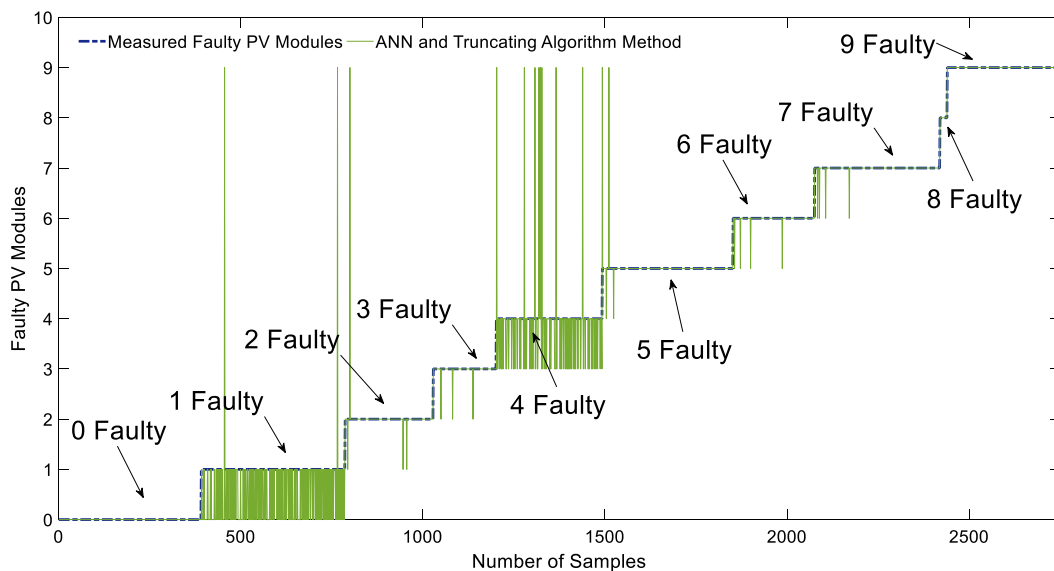


Figure 15. System 1 results using the ANN combined with a truncating algorithm

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

354 Following, we can analyse the results of the proposed method combining the ANN and fuzzy
 355 logic system. Figure 16 shows the measured faulty PV modules vs. the neuro-fuzzy fault detection

356
357

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.

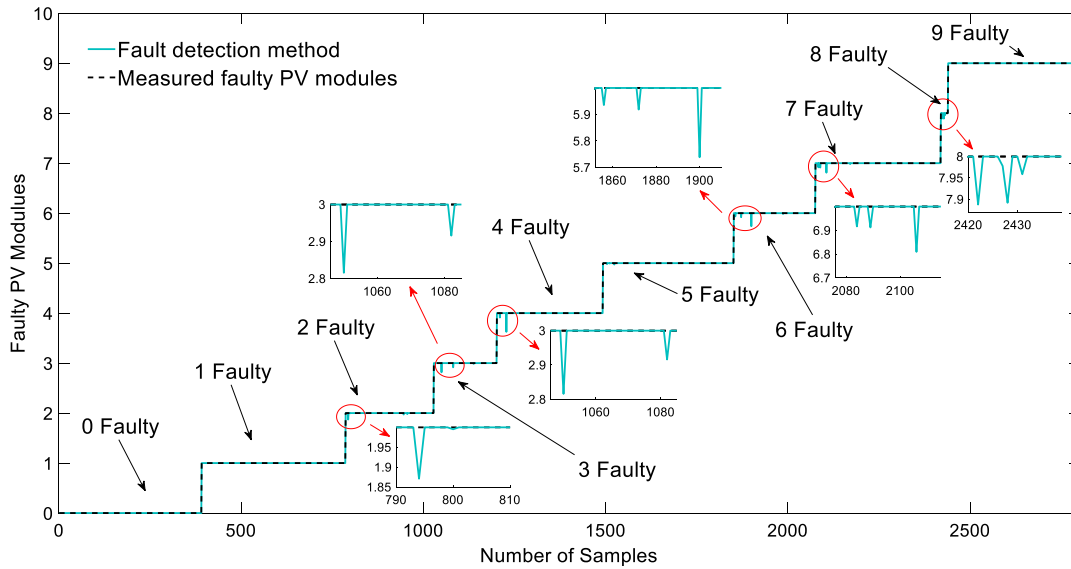


Figure 16. System 1 results using the neuro-fuzzy proposed method

358
359
360
361
362
363
364

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.

365 *5.2. System 2 Experimental Setup and Method Validation*

366
367
368
369

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.

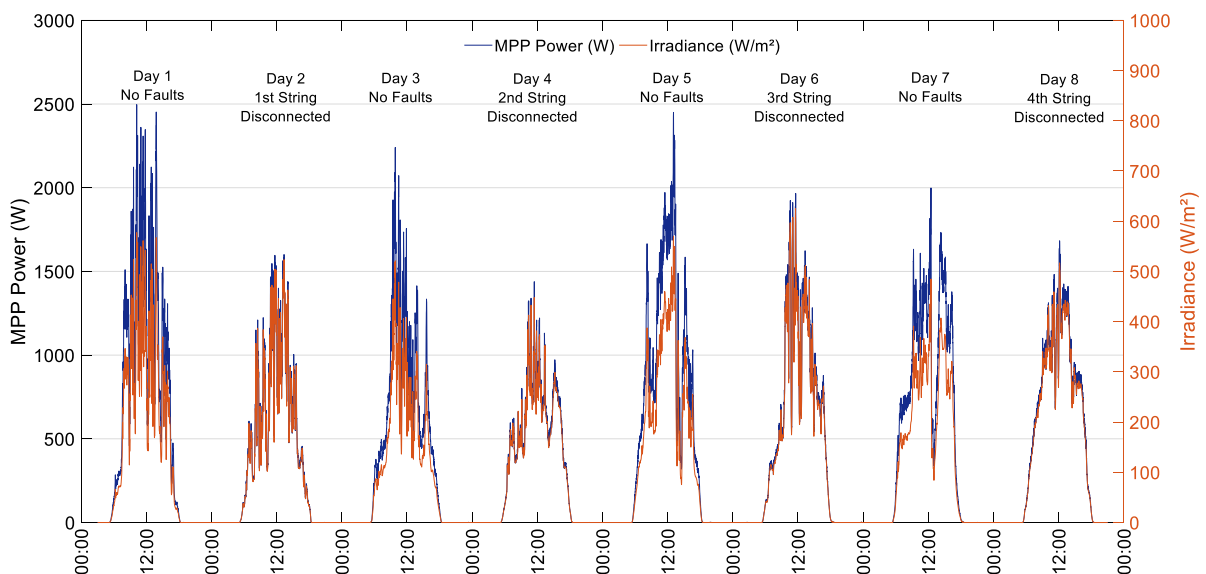


Figure 17. System 2 experimental results for irradiance G and P_{MPP}

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

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

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

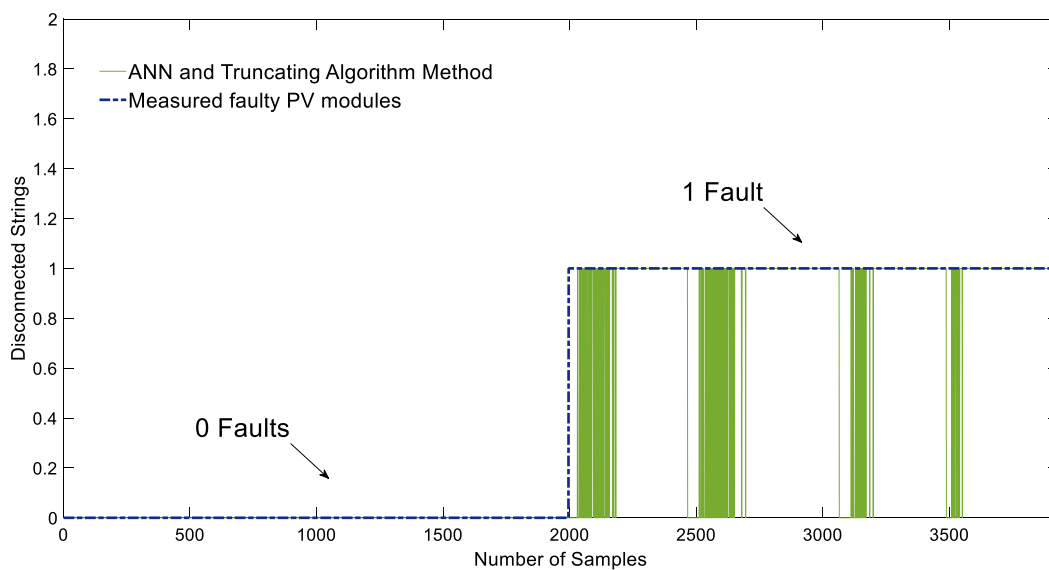


Figure 18. System 2 results using the ANN combined with a truncating algorithm

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

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

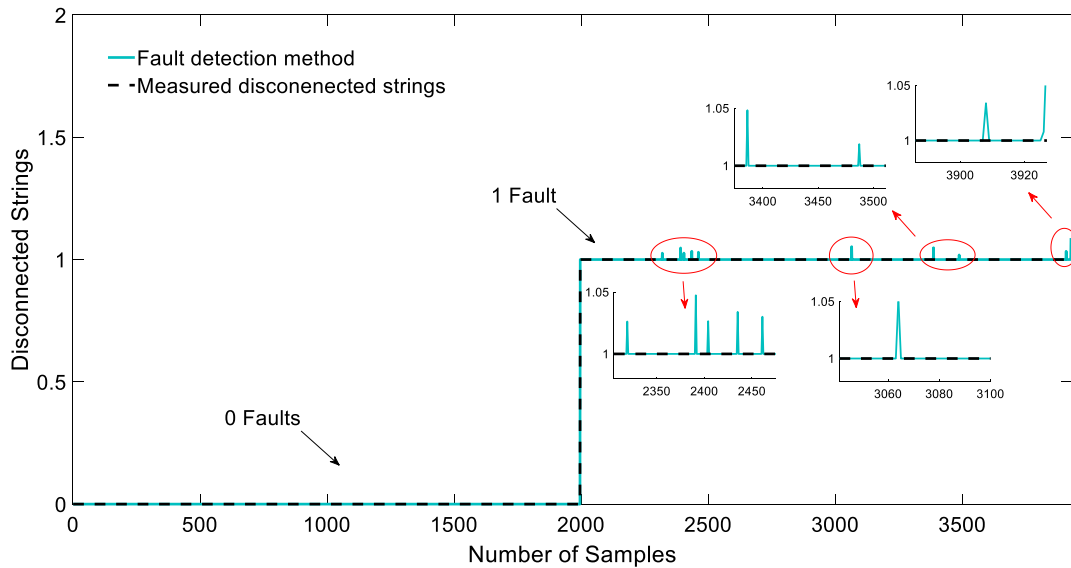


Figure 19. System 2 results using the neuro-fuzzy proposed method

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

395 6. Conclusions

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

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

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

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

418 **Funding:** This research received no external funding

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

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

423 References

- 424 1. IEA *Snapshot of Global Photovoltaic Markets*; Brussels, 2020;
- 425 2. Ghaffarzadeh, N.; Azadian, A. A Comprehensive Review and Performance Evaluation in Solar (PV)
426 Systems Fault Classification and Fault Detection Techniques. *J. Sol. Energy Res.* **2019**, *4*, 252–272.
- 427 3. Pillai, D.S.; Blaabjerg, F.; Rajasekar, N. A Comparative Evaluation of Advanced Fault Detection
428 Approaches for PV Systems. *IEEE J. Photovoltaics* **2019**, *9*, 513–527.
- 429 4. Madeti, S.R.; Singh, S.N. A Comprehensive Study on Different Types of Faults and Detection Techniques
430 for Solar Photovoltaic System. *Sol. Energy* **2017**, *158*, 161–185.
- 431 5. Dhimish, M.; Chen, Z. Novel Open-Circuit Photovoltaic Bypass Diode Fault Detection Algorithm. *IEEE*
432 *J. Photovolt.* **2019**, *9*, 1819–1827.
- 433 6. Leva, S.; Mussetta, M.; Ogliari, E. PV Module Fault Diagnosis Based on Microconverters and Day-Ahead
434 Forecast. *IEEE Trans. Ind. Electron.* **2019**, *66*, 3928–3937.
- 435 7. Syafaruddin; Karatepe, E.; Hiyama, T. Controlling of artificial neural network for fault diagnosis of
436 photovoltaic array. In *Proceedings of the 16th International Conference on Intelligent System*
437 *Applications to Power Systems*; IEEE, 25–28 September: Hersonissos, Greece, 2011; pp. 1–6.
- 438 8. Li, Z.; Wang, Y.; Zhou, D.; Wu, C. An Intelligent Method for Fault Diagnosis in Photovoltaic Array. In
439 *Proceedings of the International Conference on Electrical and Information Technologies (ICEIT)*; IEEE,
440 15–18 November: Rabat, Morocco, 2017; pp. 10–16.
- 441 9. Jiang, L.L.; Maskell, D.L. Automatic Fault Detection and Diagnosis for Photovoltaic Systems Using
442 Combined Artificial Neural Network and Analytical Based Methods. In *Proceedings of the Proceedings*
443 *of the International Joint Conference on Neural Networks*; IEEE, 12–17 July: Killarney, Ireland, 2015; Vol.
444 2015-Septe.
- 445 10. Akram, M.N.; Lotfifard, S. Modeling and Health Monitoring of DC Side of Photovoltaic Array. *IEEE*
446 *Trans. Sustain. Energy* **2015**, 1–9.
- 447 11. Garoudja, E.; Chouder, A.; Kara, K.; Silvestre, S. An Enhanced Machine Learning Based Approach for
448 Failures Detection and Diagnosis of PV Systems. *Energy Convers. Manag.* **2017**, *151*, 496–513.
- 449 12. Chine, W.; Mellit, A.; Lughy, V.; Malek, A.; Sulligoi, G.; Massi Pavan, A. A Novel Fault Diagnosis
450 Technique for Photovoltaic Systems Based on Artificial Neural Networks. *Renew. Energy* **2016**, *90*, 501–
451 512.
- 452 13. Dhimish, M.; Holmes, V.; Mehrdadi, B.; Dales, M.; Mather, P. Photovoltaic Fault Detection Algorithm
453 Based on Theoretical Curves Modelling and Fuzzy Classification System. *Energy* **2017**, *140*, 276–290.
- 454 14. Dhimish, M.; Holmes, V.; Mehrdadi, B.; Dales, M. Multi-Layer Photovoltaic Fault Detection Algorithm.
455 *High Volt.* **2017**, *2*, 244–252.
- 456 15. Dhimish, M.; Holmes, V.; Mehrdadi, B.; Dales, M. Comparing Mamdani Sugeno fuzzy logic and RBF
457 ANN network for PV fault detection. *Renew. Energy* **2018**, *117*, 257–274.
- 458 16. Hussain, M.; Dhimish, M.; Titarenko, S.; Mather, P. Artificial Neural Network Based Photovoltaic Fault
459 Detection Algorithm Integrating Two Bi-Directional Input Parameters. *Renew. Energy* **2020**, *155*, 1272–
460 1292.
- 461 17. Benghanem, M.S.; Alamri, S.N. Modeling of Photovoltaic Module and Experimental Determination of
462 Serial Resistance. *J. Taibah Univ. Sci.* **2009**, *2*, 94–105.
- 463 18. Villalva, M.G.; Gazoli, J.R.; Filho, and E.R. Comprehensive Approach to Modeling and Simulation of

- 464 Photovoltaic Arrays. *IEEE Trans. POWER Electron.* **2009**, *24*, 1198–1208.
- 465 19. Khatib, T.; Elmenreich, W. *Modeling of Photovoltaic Systems Using MATLAB*; 1st ed.; John Wiley & Sons:
466 Hoboken, 2016; Vol. 3; ISBN 9781119118107.
- 467 20. Bayhan, H.; Bayhan, M. A Simple Approach to Determine the Solar Cell Diode Ideality Factor Under
468 Illumination. *Sol. Energy* **2011**, *85*, 769–775.
- 469 21. Meyer, E.L. Extraction of Saturation Current and Ideality Factor from Measuring Voc and Isc of
470 Photovoltaic Modules. *Int. J. Photoenergy* **2017**, 2017.
- 471 22. Siddique, N.; Adeli, H. *Computational Intelligence Computational Intelligence Synergies of Fuzzy Logic, Neural
472 Networks and Evolutionary Computing*; 1st ed.; John Wiley & Sons: New Delhi, India, 2013; ISBN
473 9781118337844.
- 474 23. Dhimish, M.; Badran, G. Photovoltaic Hot-Spots Fault Detection Algorithm Using Fuzzy Systems. *IEEE
475 Trans. Device Mater. Reliab.* **2019**, *19*, 671–679.
- 476 24. Engelbrecht, A.P. *Computational intelligence: An introduction*; 2nd ed.; Jonh Wiley & Sons, Ltd: Chichester,
477 West Sussex, United Kingdom, 2007; Vol. 1;
- 478 25. Dhimish, B.M.; Holmes, V.; Mehrdadi, B.; Dales, M.; Mather, P. Detecting Defective Bypass Diodes in
479 Photovoltaic Modules using Mamdani Fuzzy Logic System. *Glob. J. Res. Eng. Electr. Electron. Eng.* 2017,
480 17, 33–42.



© 2020 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).