



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/185918/>

Version: Accepted Version

Article:

Vieira, Romênia G., Dhimish, Mahmoud, M. U. de Araújo, Fábio et al. (2022) Comparing multilayer perceptron and probabilistic neural network for PV systems fault detection. Expert systems with applications. 117248. ISSN: 0957-4174

<https://doi.org/10.1016/j.eswa.2022.117248>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

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.

Comparing Multilayer Perceptron and Probabilistic Neural Network for PV Systems Fault Detection

Romênia Gurgel Vieira^{1,*}, Mahmoud Dhimish², Fábio Meneghetti Ugulino de Araújo³, Maria Izabel da Silva Guerra¹

¹*Department of Engineering and Technology, Semi-Arid Federal University, Francisco Mota Av., Mossoro, 59625-900, Brazil;*

²*Department of Electronic Engineering, University of York, York YO10 5DD, United Kingdom;*

³*Department of Computer and Automation Engineering, Federal University of Rio Grande do Norte, Natal, 59078-970, Brazil;*

*Corresponding author.

Email addresses: romenia.vieira@ufersa.edu.br (Romênia Gurgel Vieira), mahmoud.dhimish@york.ac.uk (Mahmoud Dhimish), meneghet@dca.ufrn.br (Fábio Meneghetti Ugulino de Araújo), izabel.guerra@ufersa.edu.br (Maria Izabel da Silva Guerra)

1 **Abstract** – This work introduces the development of a fault detection method for photovoltaic (PV)
2 systems using artificial neural networks (ANN). The faults identified by the method are short-circuited
3 modules and disconnected strings. This research's novel part is its adaptability as a long-term dataset
4 has been used in the ANN training and validation phase and also examined situations considering
5 datasets contaminated with random noise. It makes the method suitable for any photovoltaic power
6 plant, also does not require long datasets from pre-existing systems or installing new sensors. The
7 proposed method comprises two unique algorithms for PV fault detection, a Multilayer Perceptron,
8 and a Probabilistic Neural Network. The research method used modeling, simulation, and experiment
9 data since both algorithms were trained using simulated datasets and tested through experimental
10 data from two different photovoltaic systems. Even though the training dataset includes noisy
11 situations, the results indicated a superior precision for the Multilayer Perceptron neural network.
12 The findings showed a maximum accuracy of 99.1% in detecting short-circuited modules and 100% in
13 detecting disconnected strings.

14 **Keywords:** Solar Energy; Photovoltaic Modules; String Disconnection; Short-circuit; Fault
15 Detection; Neural Network.

Nomenclature

AC	Alternate Current
ANN	Artificial Neural Network
DC	Direct Current
MLP	Multilayer Perceptron
MPP	Maximum Power Point
MPPT	Maximum Power Point Tracking
NOCT	Nominal Operating Cell Temperature
PDF	Probability Density Functions
PNN	Probabilistic Neural Network
PV	Photovoltaic
P-V	Power versus Voltage
RBF	Radial Basis Function
ROC	Receiver Operating Characteristics

1. Introduction

17 Solar photovoltaic (PV) technology has been introduced as a renewable source of energy
18 worldwide. It is not only a clean choice but also free and available. PV systems across the world
19 reached a total installed capacity of 627 GW (IEA, 2020). Moreover, PV technology shows significant
20 flexibility, considering it can be incorporated into constructions, comprehending industrial,
21 commercial and domestic buildings.

22 PV systems are subject to several fault conditions during their operation. Such conditions may
23 impact the system's reliability, decreasing its performance and lifetime, and in some cases leaving the
24 whole operation in danger. Faults in PV systems can occur on the DC or AC side, affecting the PV
25 modules, converters, maximum power point trackers (MPPT), or inverters. Some of these faults can
26 be hard to detect, decreasing the power production for long periods. Faults arising on PV systems may
27 reduce the generation by 18.9% (Pillai, Blaabjerg, & Rajasekar, 2019).

28 The PV modules are the primary generation unit, so faults occurring on such devices profoundly
29 impact the PV system's reliability. Such faults can be permanent or temporary, depending on their
30 source (Madeti & Singh, 2017). There are various causes of PV module faults, like mismatch faults,
31 bypass diodes (Vieira, de Araújo, Dhimish, & Guerra, 2020), module aging, potential induced
32 degradation (Dhimish, Hu, Schofield, & Vieira, 2020), shading, short-circuit faults, and string
33 disconnections.

34 Therefore, quickly detecting and diagnosing PV systems' faults is crucial for reliability and
35 avoiding high maintenance costs. Accordingly, this section discusses the research background in the
36 field, especially regarding the machine learn-based fault detection methods, as discuss our
37 contributions to knowledge.

1.1. Literature Review

38 In recent years, several fault detection methods have been studied. It can be classified into two
39 groups: electrical and nonelectrical methods. Among the electrical methods, it is found statistical
40 methods, signal processing, and machine learning techniques (Ghaffarzadeh & Azadian, 2019).

41 Regarding the machine learning methods, Syafaruddin *et al.* (2011) proposed a feedforward
42 artificial neural network (ANN) for detecting and localizing short-circuit PV modules. The authors
43 used module temperature, irradiance and current, and voltage at the maximum power point (MPP) as
44 input variables. The method was tested on a six-module array, showing promising results.

45 Another ANN using the same input variables as Syafaruddin *et al.* (2011) was studied by Li *et al.*
46 (2017). This method identifies and localizes short-circuited PV modules, degradation, and shading
47 faults. They extracted the training dataset using MATLAB/Simulink[®] simulations, and the algorithm

48 Was not experimentally tested. Jiang and Maskell (2015) proposed an ANN combined with an
 49 analytical method. The ANN predicts the expected MPP using temperature and irradiance as input
 50 variables. The analytical algorithm compares the ANN result to the measured MPP, enabling the
 51 diagnosis of open-circuited string or module, short-circuited module, partial shading, and
 52 malfunctioning at the MPPT unit. This method was not experimentally tested.

53 A short-circuit and open-circuit fault detection method developed by Akram and Lotfifard (2015)
 54 applies a probabilistic neural network (PNN). The training dataset was compiled by simulations using
 55 MATLAB/Simulink® software. The authors tested the algorithm also using simulated data, showing a
 56 maximum error of 3.5%. Later, Garoudja *et al.* (2017) also applied a PNN for detecting short-circuited
 57 PV modules and disconnected strings. The input variables are temperature, irradiance, voltage, and
 58 current at the MPP, and the training dataset is extracted by simulation. This method was experimentally
 59 tested, and as a result, the authors compared the PNN performance to an ANN. The proposed PNN
 60 showed 100% accuracy in detecting the approached faults, while the ANN showed 90.3%.

61 One more ANN fault detection method was developed by Dhimish *et al.* (2018). The authors
 62 compare a fuzzy logic system to a radial basis function (RBF) network for detecting partial shading,
 63 short-circuited PV module, and MPPT malfunctioning. The results showed an accuracy of 92.1% for
 64 the RBF algorithm, superior then the fuzzy logic.

65 Vieira *et al.* (2020b) proposed a fault detection technique combining ANN and fuzzy logic system.
 66 The method diagnoses short-circuited and disconnected strings on a PV system using input variables,
 67 ambient temperature, irradiance, and power at the MPP. The authors validated the method using
 68 experimental data, showing an accuracy of 99.43% for detecting short-circuited PV modules and
 69 99.43% for disconnected strings.

1.2. Related Studies

70 Considering the extensive discussion, several studies explored fault detection methods. However,
 71 as we can observe from Table 1, most of them require data from pre-existing systems, installing extra
 72 sensors on the PV plant and some of its methodologies need to compare simulated results to measured
 73 data, *i.e.*, uses a residual error or a rate to indicate the presence of a fault, which makes the process
 74 more complex. Also, it is essential to highlight that none of the explored researches considered noisy
 75 situations for the training data.

Table 1 – Discussed fault detection methods

Reference	Experimentally tested	Training Dataset	Extra sensors	Residual Error/Rate	Noisy Situation
(Syafaruddin et al., 2011)	No	Simulated	Yes	Yes	No
(Li, Wang, Zhou, & Wu, 2012)	No	Simulated	Yes	No	No
(Jiang & Maskell, 2015)	No	Simulated	No	Yes	No

(Akram & Lotfifard, 2015)	No	Simulated	No	No	No
(Chine et al., 2016)	Yes	Experimental	Yes	Yes	No
(Garoudja et al., 2017)	Yes	Simulated	No	No	No
(Madeti & Singh, 2018)	No	Simulated	No	Yes	No
(Dhimish et al., 2018)	Yes	Experimental	No	Yes	No
(Vieira, Dhimish, et al., 2020)	Yes	Simulated	No	No	No

76 Table 1 and the previous section demonstrate a lack of research experimental results on fault
77 detection methods, and mainly that those studies do not investigate noisy situations. Therefore, this
78 paper proposes and compares two fault detection techniques using different neural networks: MLP
79 (Multilayer Perceptron) and PNN (Probabilistic Neural Network). The main contribution of this
80 research is to develop an algorithm capable of detecting faults on PV systems and analyzing their
81 performance under noisy situations. The faults detected by the algorithms are short-circuited PV
82 modules and string disconnections. These faults, as earlier described, can reduce the generated PV
83 power, and observing it can be costly and time-consuming. The proposed method does not require a
84 long-term dataset from pre-existing PV systems, installing extra sensors, and was experimentally
85 tested.

86 The paper is briefly structured as follows. Section 2 defines the methodology used to develop the
87 short-circuited PV modules detection method, presenting the studied PV systems and the experimental
88 setup for testing the proposed fault detection methods. Then, Section 2.2 presents the proposed
89 algorithms' results and discussion, analyzing their performance with experimental data of the studied
90 PV systems. Finally, in Section 4, the overall conclusions are discussed.

2. Research Methodology

91 To develop the proposed research, we followed five stages, as illustrated in Fig. 1. At first, we
92 modeled and simulated the studied PV systems using MATLAB/Simulink[®]. Then, we validated the
93 developed simulation with experimental data.

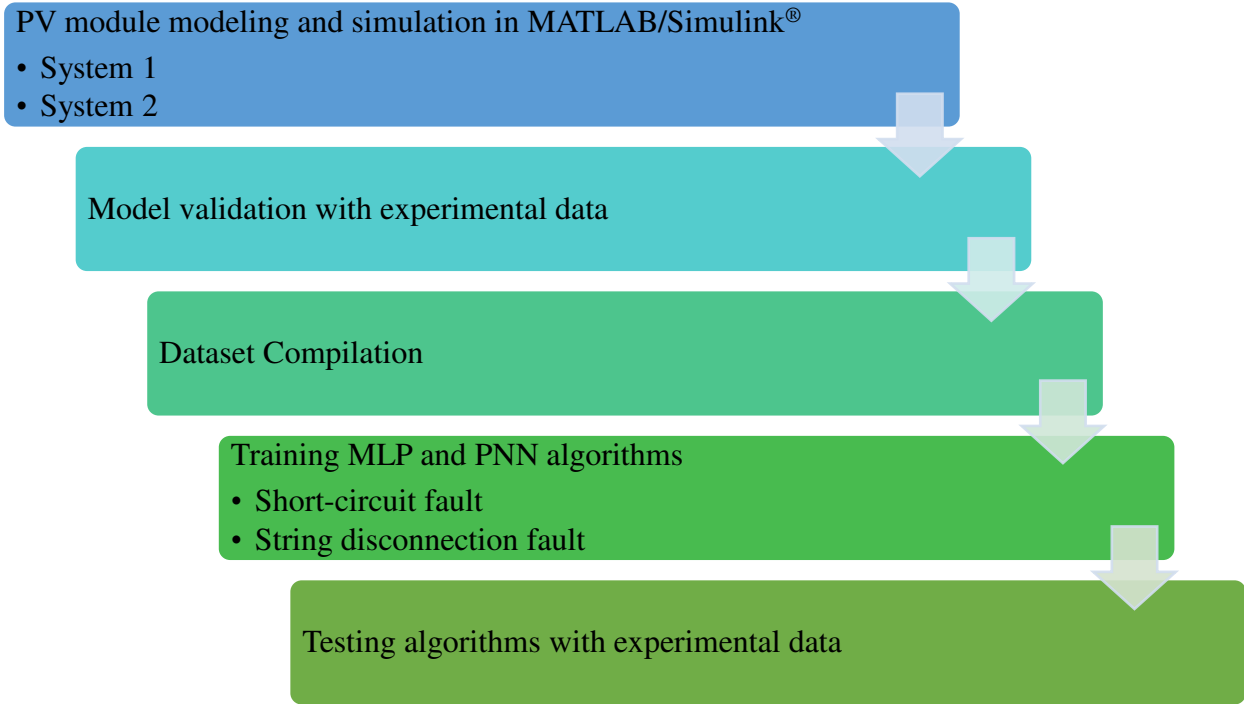


Fig. 1 - Research methodology workflow

94 Since the model was validated, it was possible to build the dataset to train the proposed fault
 95 detection algorithms. After training the neural networks to detect short-circuited PV modules and
 96 string disconnection faults, we tested the proposed method using experimental data to assess its
 97 accuracy in detecting faults on PV systems, approaching all studied scenarios.

2.1. System 1 and 2: Description and Model Validation

98 The PV module model employed in this research is based on the one diode model, considering its
 99 simplicity. The model simulation was developed and extensively discussed in a previous work
 100 published by the authors (Guerra, Ara, Dhimish, & Vieira, 2021; Vieira, Dhimish, et al., 2020).

101 We examined two different PV systems, named here as System 1 and System 2. Both power plants
 102 were experimentally tested to validate the model simulation and the proposed fault detection methods.

103 The first studied PV plant is a 2.2 kWp system installed at the Huddersfield University campus. It
 104 consists of one string with ten series-connected PV modules. The modules model is the SMT6(60)P
 105 from PowerGlaz manufacturer, with a nominal power of 220 W (per module). Table 2 describes the
 106 PV modules' electrical parameters.

Table 2 - System 1 PV module characteristics

Datasheet parameters			
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			
R_{sh}	1108.3972 Ω	R_s	0.3930 Ω

107 We experimentally tested System 1 under healthy and faulty conditions. The conducted tests
 108 disconnected the PV modules using the connection box (see Fig. 2) to emulate the short-circuit fault
 109 condition. Therefore, we created ten scenarios, the first one with no faulty conditions, followed by 1,
 110 2, 3 until 9 faulty conditions, as illustrated in Fig. 2.

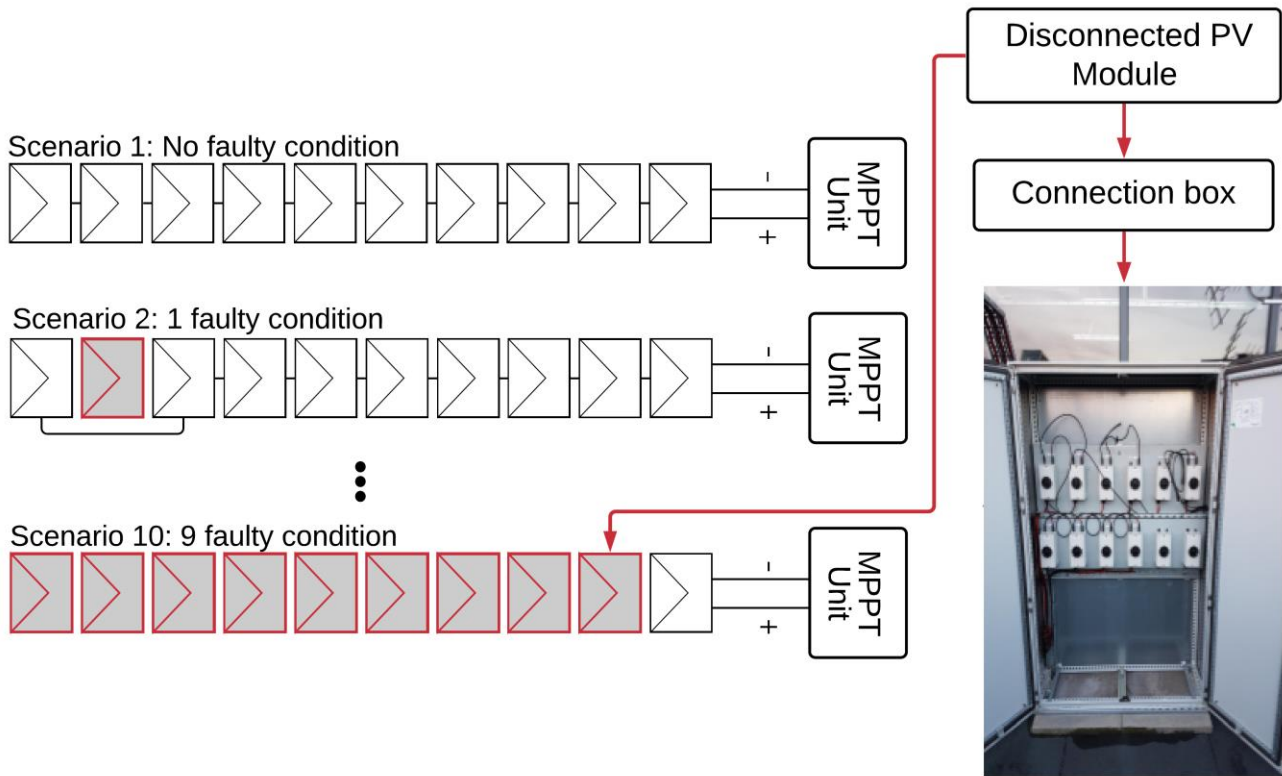


Fig. 2 - System 1 experimental setup

111 We performed the experiments for two weeks, observing each faulty scenario for the whole day.
 112 During the tests, we measured the peak power (P_{MPP}) as an electrical variable and the irradiance (G)
 113 and ambient temperature (T_a) as nonelectrical variables. The measured temperature was constant
 114 through the observed days, approximately 16 °C, and the results for P_{MPP} and G are illustrated in Fig.
 115 3 and Fig. 4.

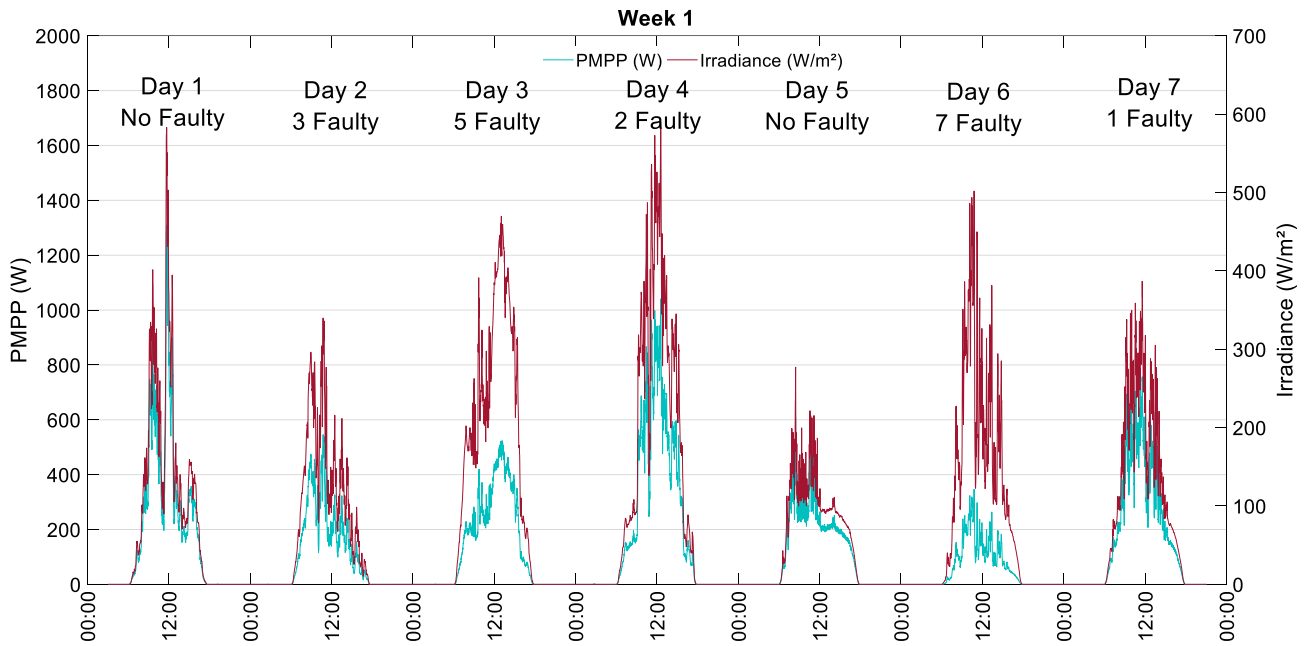


Fig. 3 - Week 1 experimental results for System 1

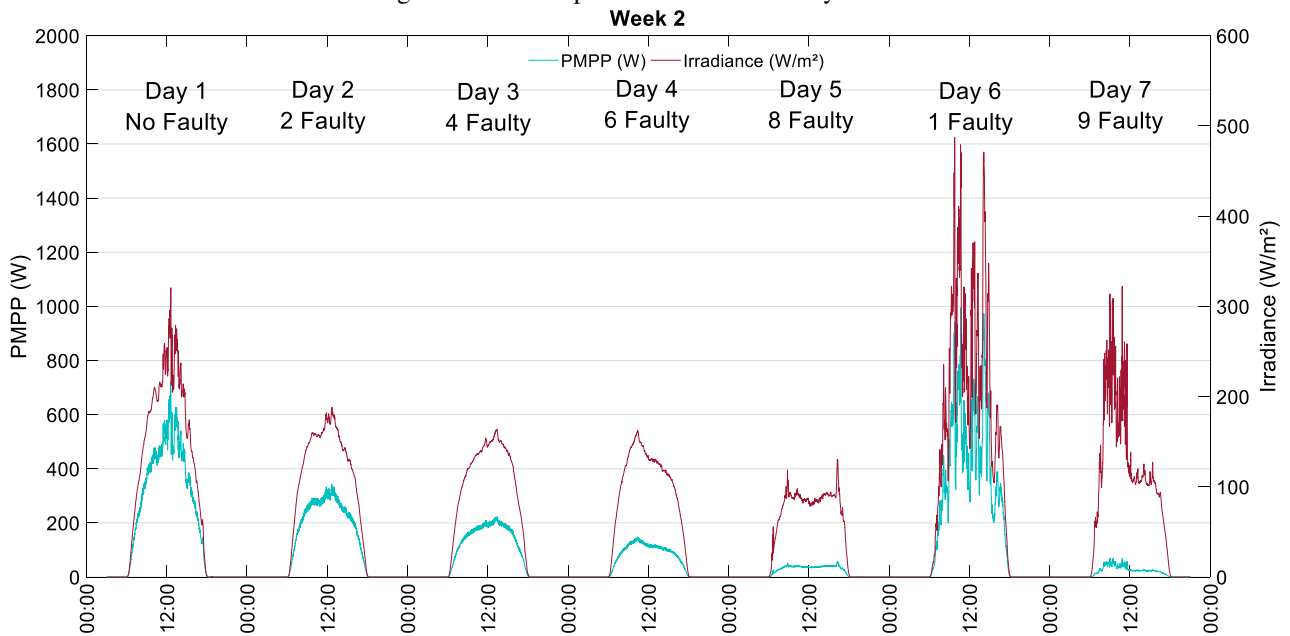


Fig. 4 - Week 2 experimental results for System 1

116 Observing Fig. 3 and Fig. 4, the P_{MPP} decreases drastically when the faulty condition arises. If we
 117 compare a typical operation day, like Day 1 in Fig. 3, to a faulty condition day, like Day 7 in Fig. 4,
 118 we can observe that as the irradiance increases, the P_{MPP} does not follow it, emphasizing that the
 119 irradiance increases the faulty condition.

120 The diode ideality factor (n) used in the model simulation was empirically chosen as 1 to improve
 121 the model fitting. We simulated System 1 using the proposed one diode model (Guerra et al., 2021;
 122 Vieira, Dhimish, et al., 2020) and compared it to experimental data from the studied system. The
 123 outcomes are described in Table 3.

124

Table 3 – System 1 modeling validation

T_a (°C)	G (W/m ²)	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

125 We can observe from Table 3 that the error between the simulation and the measured data is
 126 minimum. Thus, we can build the training dataset using this model simulation for System 1.

127 The second studied PV system is a 4.16 kWp power plant also installed at the Huddersfield
 128 University campus. It comprises 32 PV modules arranged into four strings, with eight modules each.
 129 The module model is the KC130GHT-2 from Kyocera manufacturer, with a nominal power of 130 W,
 130 and its electrical characteristics are described in Table 4.

Table 4 - System 2 PV module characteristics

Datasheet parameters			
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			
R_{sh}	119.232 Ω	R_s	0.16 Ω

131 System 2 was also experimentally tested under healthy and faulty conditions. In this case, we
 132 disconnected the strings one at a time, starting with the first string, followed by the second, third, and
 133 fourth, to emulate the string disconnection faulty condition. Then, the strings were disconnected using
 134 the switch box, as Fig. 5 illustrates.

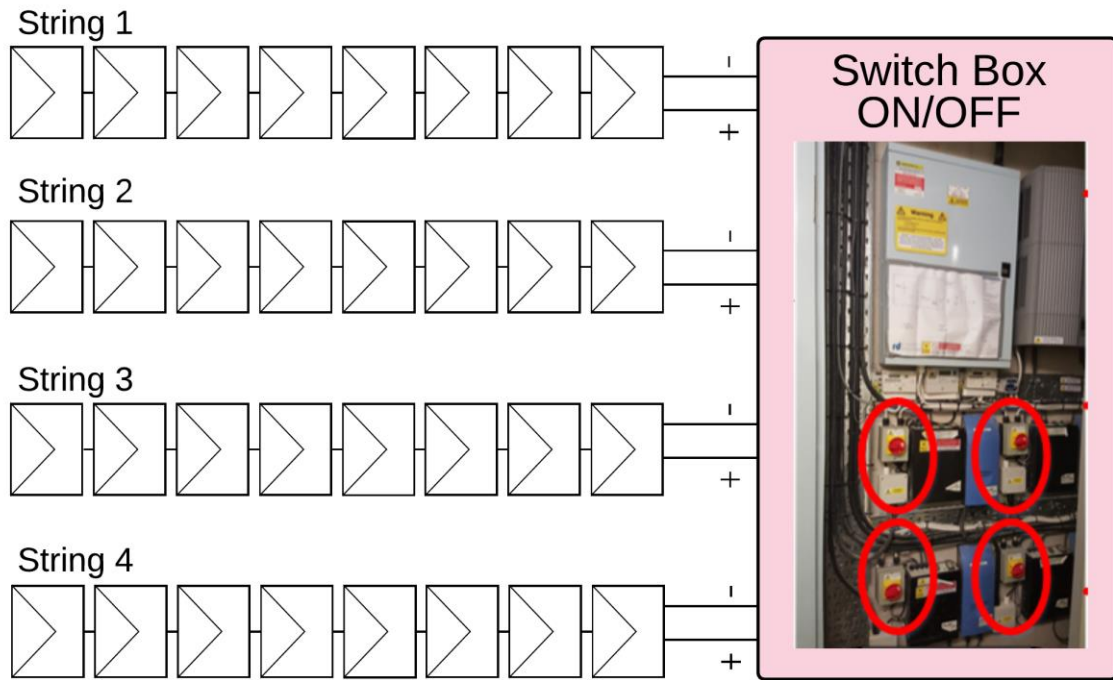


Fig. 5 - System 2 experimental setup

135 In System 2, we performed the tests for eight days, observing each faulty condition for the
 136 whole day. The measured variables were also peak power (P_{MPP}), irradiance (G), and ambient
 137 temperature (T_a). The ambient temperature was around 16 °C, and the experimental results for P_{MPP}
 138 and G are illustrated in Fig. 6.

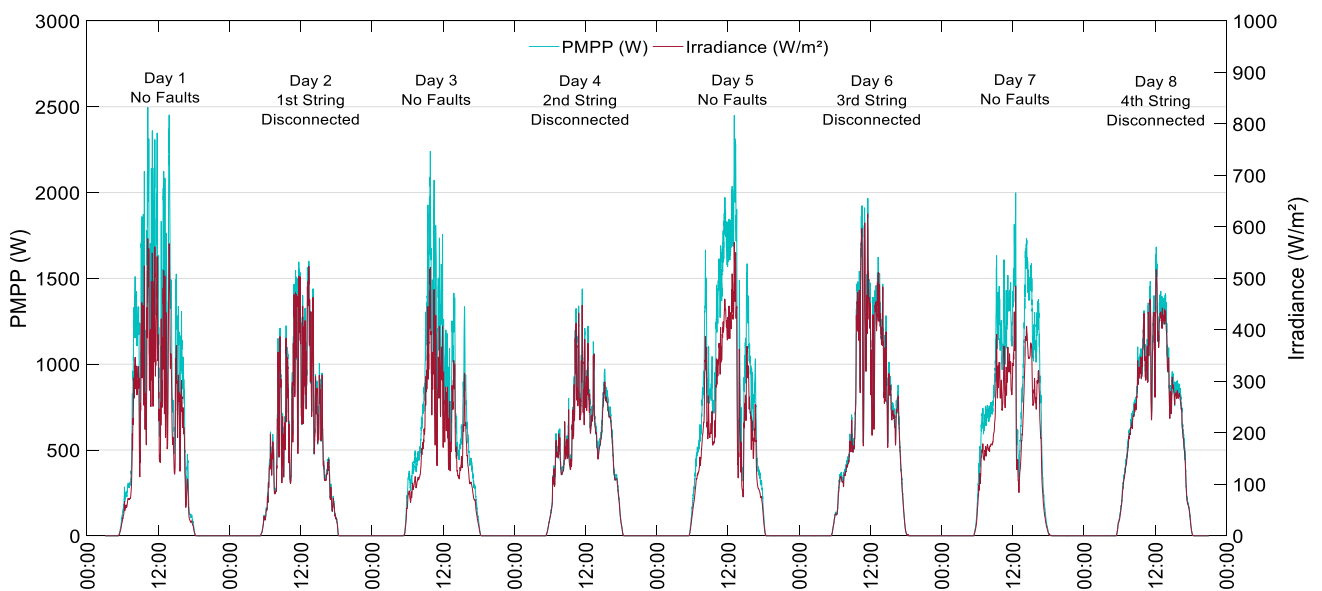


Fig. 6 - Experimental results for System 2

139 Observing Fig. 6, we note that the peak power (P_{MPP}) decreases when the system operates under
 140 faulty conditions. Comparing Day 1 (no faults) to Day 5 (one string disconnected), the output power
 141 does not increase as the irradiance increase. This situation underlined the faulty condition occurring
 142 on System 2.

143 The diode ideality factor (n) used in the model simulation was empirically chosen as 1.2 to
 144 improve the model fitting. We also simulated System 2 using the proposed one diode model (Guerra
 145 et al., 2021; Vieira, Dhimish, et al., 2020) and compared it to experimental data from the studied
 146 system. The comparison between simulated and measured data is described in Table 5.

Table 5 - System 2 modeling validation

T_a (°C)	G (W/m ²)	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

147 We can observe from Table 5 that the error between the simulation and the measured data is
 148 minimum, enabling us to assemble the required training dataset by simulation.

149

2.2. Detecting Short-Circuit PV Modules: MLP and PNN

150 This Section describes the neural networks developed for detecting short-circuit PV modules on
 151 System 1. We extracted the training dataset using the authors' previous model simulation (Vieira,
 152 Dhimish, et al., 2020).

153 We developed the algorithms for the short-circuited PV module faulty condition considering
 154 System 1. The obtained dataset comprises 7070 samples, 707 for each faulty condition. We settled
 155 three scenarios for evaluating the algorithms: Scenario 1, Scenario 2, and Scenario 3. The first case,
 156 Scenario 1, corresponds to the raw data extracted by simulation. For the others examined conditions,
 157 we inserted a noise of $\pm 15\%$ in the P_{MPP} input variable. Thus, Scenario 1 is a noiseless condition, while
 158 Scenario 2 contains a noise of $\pm 15\%$ on 50% of the MPP data, and Scenario 3 contains the noise in
 159 100% of the MPP data.

160 This noise represents the uncertainties associated with sensors, amplifiers, and analog and digital
 161 converters, resulting in incorrect measurements and tricks the MPPT algorithm into settling on the
 162 incorrect MPP (Al-Atrash, Batarseh, & Rustom, 2010). Therefore, we can evaluate how the algorithms
 163 respond when trained with noisy data.

164 Thus, the research offers two neural network types, MLP and PNN, to compare and analyze which
 165 neural network is more suitable to tackle this faulty problem, considering each specified scenario, as
 166 illustrated in the scheme in Fig. 7.

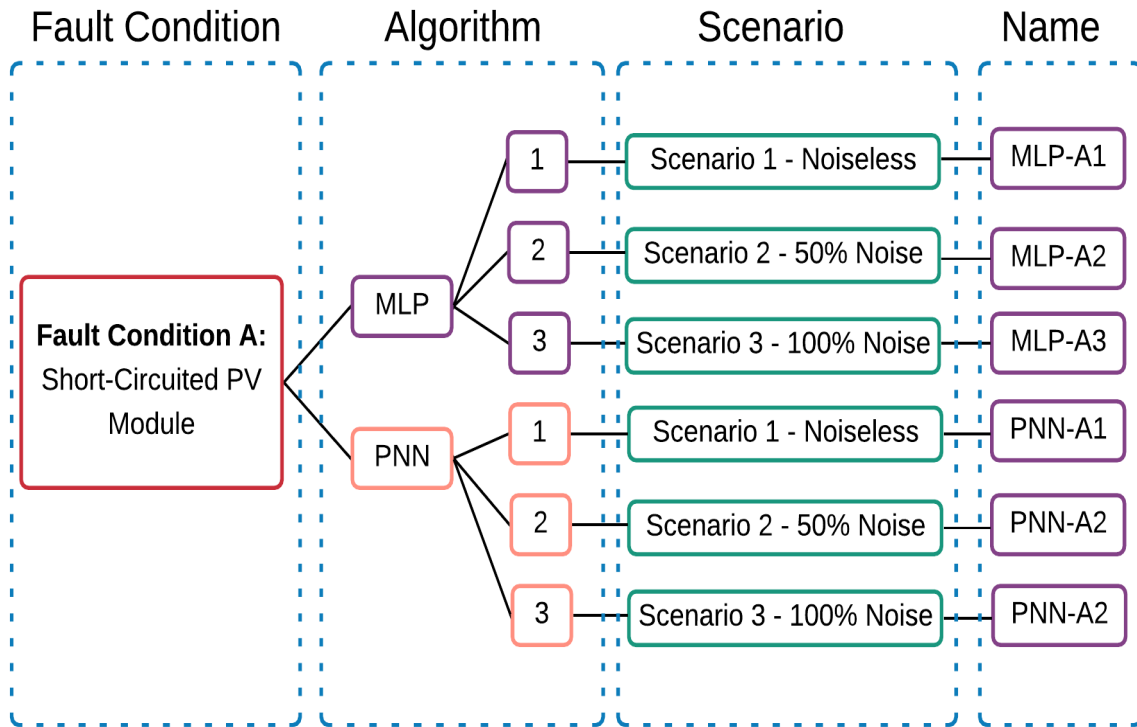


Fig. 7 - Schematic of the studied algorithms and conditions for detecting short-circuited PV modules

167 The first algorithm employed as a fault detection method is a multilayer perceptron (MLP) neural
 168 network. MLP neural nets are characterized by the presence of at least one hidden layer and an output
 169 layer. The signal flow starts at the input layer, then passes through the intermediate layer, and ends at
 170 the output neural layer, as illustrated in Fig. 8.

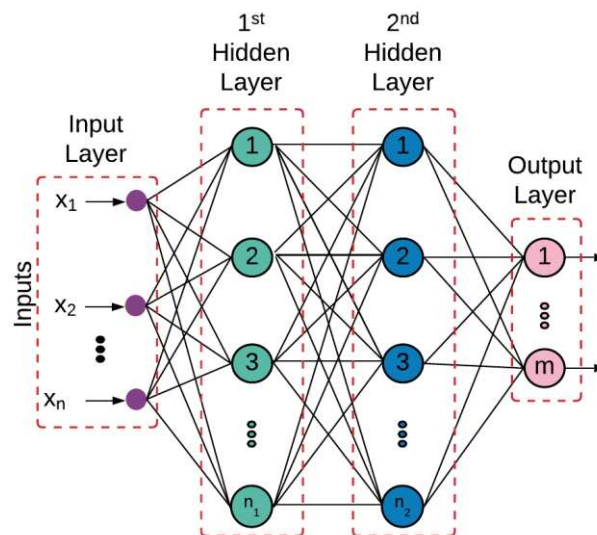


Fig. 8 - MLP basic structure

171 Generally, MLP networks are employed in various situations since pattern recognition, process
 172 identification and control, and systems optimization. There are no strict guidelines on deciding the

173 number of neurons and hidden layers, although it influences network performance. For instance, many
 174 neurons in the hidden layer can produce better results and make the training process low (Siddique and
 175 Adeli, 2013).

176 We trained three networks, one for each studied scenario, using the same structure in all cases.
 177 Fig. 9 shows the neural nets' structure developed using MATLAB[®] software, and Table 6 describes its
 178 training settings.

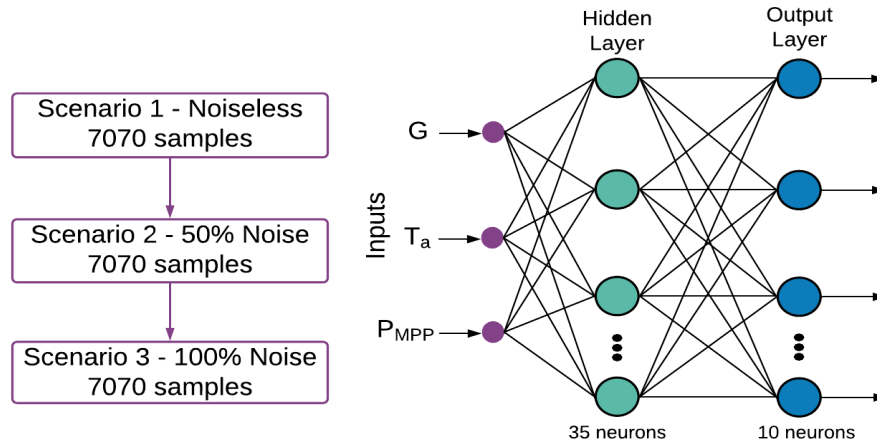


Fig. 9 - MLP network structure for detecting short-circuit PV modules

Table 6 - MLP training characteristics for short-circuited PV modules detection

MLP	
Input Variables	3 (G, T, P _{MPP})
Output Variables	10
Number of Layers	3
Number of Neurons	(35, 10)
Training Process	supervised
Training Algorithm	Levenberg-Marquardt
Activation Function	(tansingmoid, tansingmoid)
Training	70%
Validation	15%
Test	15%
Type of Divison Samples	random

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

184 The input variables are irradiance (G), ambient temperature (T_a), and the maximum power point
 185 (P_{MPP}). The output is a vector equals zero, except for one element equals 1. This element represents

186 the faulty condition identified. For System 1, there are ten faulty classes. The first one represents
 187 normal operation. Table 7 represents the output vectors for the trained MLPs.

Table 7 - Output vectors for System 1 MLPs

Short-circuited modules	Fault	Output	Class
Normal Operation	F0	1 0 0 0 0 0 0 0 0 0	1
1	F1	0 1 0 0 0 0 0 0 0 0	2
2	F2	0 0 1 0 0 0 0 0 0 0	3
3	F3	0 0 0 1 0 0 0 0 0 0	4
4	F4	0 0 0 0 1 0 0 0 0 0	5
5	F5	0 0 0 0 0 1 0 0 0 0	6
6	F6	0 0 0 0 0 0 1 0 0 0	7
7	F7	0 0 0 0 0 0 0 1 0 0	8
8	F8	0 0 0 0 0 0 0 0 1 0	9
9	F9	0 0 0 0 0 0 0 0 0 1	10

188 Table 8 describes the outcomes for the MLPs network training process. Observing the training
 189 accuracy results, the MLP-A1 showed an accuracy of 99.9% on training, while MLP-A2 and A3
 190 showed 85.5% and 70.4%, respectively. So, we notice that the training accuracy drastically decreases
 191 when we insert the noise on the dataset.

Table 8 - MLPs training results for detecting short-circuit PV modules

MLP	Epochs	Regression Coefficient	Training Accuracy
A1	69	0.99878	99.9%
A2	74	0.86846	85.5%
A3	52	0.78320	70.4%

192 The next algorithm tested on this research is a Probabilistic Neural Network (PNN). PNN's
 193 neural networks are feedforward neural nets based on statistical principles instead of heuristic methods.
 194 In general, heuristic approaches continuously modify the algorithm's parameters to improve network
 195 performance gradually. The MLP is an example of a heuristic method that requires long training but
 196 does not always reach the best solution within a reasonable time (Siddique and Adeli, 2013).

197 A PNN network is a simple parallel three-layer derived from Bayes decision strategy and
 198 nonparametric kernel-based estimators of probability density functions (PDF). The most common
 199 PNN method uses the sum of spherical Gaussian functions centered at each training vector to estimate
 200 the PDFs' class. Equation (1) and Fig. 10 describe a PNN network's basis (Siddique and Adeli, 2013).

$$f_i(x) = \frac{1}{(2\pi)^{\frac{p}{2}} \sigma^p M} \times \frac{1}{M} \sum_{j=1}^M \exp \left[\frac{-(x - x_{ij})^T (x - x_{ij})}{2\sigma^2} \right] \quad (1)$$

201 Where i represents the class number, and j the pattern number, x_{ij} is the j^{th} training vector from i ,
 202 x is the test vector, M is the number of test vectors in i , p is the dimension of the vector x , σ is the

203 smoothing factor and $f_i(x)$ is the sum of multivariate Gaussian distribution centered at each of the
 204 training samples.

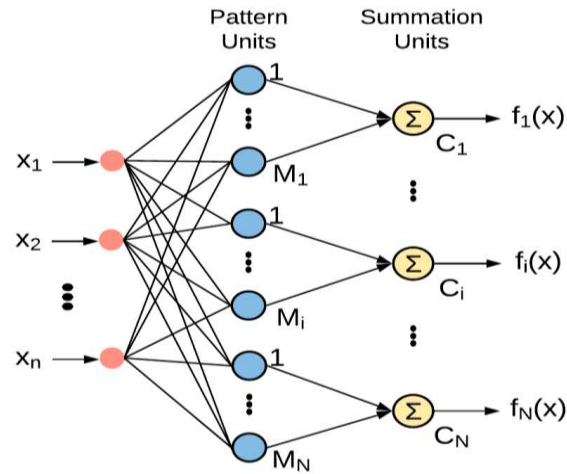


Fig. 10 - PNN basic structure

205 Training a PNN network is fast and easy. However, it requires lots of memory space, considering
 206 that all training vectors must be stored and used (Siddique & Adeli, 2013). Therefore, analogous to the
 207 MLPs network, we trained three networks, one for each studied scenario, using the same structure in
 208 all cases. Fig. 11 shows the neural nets' structure developed using MATLAB[®] software.

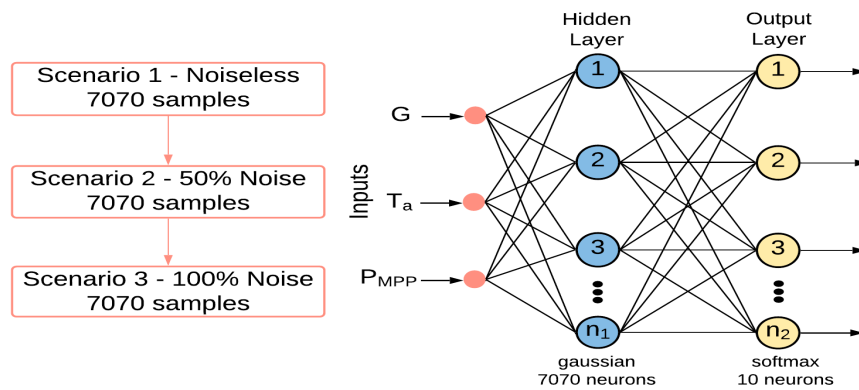


Fig. 11 - PNN network structure for detecting short-circuit PV modules

209 The input and output variables are equal to those used for the MLPs networks, following the same
 210 output vector described in Table 7. The hidden and output layers activation functions are gaussian and
 211 softmax, respectively.

2.3. Detecting Disconnected Strings: MLP and PNN

212 This Section describes the neural networks developed for detecting disconnected strings on
 213 System 2. The fault detection methods used a simulated dataset, analog to the procedure described for
 214 detecting short-circuited PV modules (see Section 2.2). However, for the string disconnection fault
 215 condition, we used System 2, described in Section 2.1.

216 The training dataset comprises 2828 samples, 707 for each faulty scenario. Just like we proceeded
 217 for the short-circuited PV modules fault condition, we examined the same three scenarios for the
 218 proposed algorithms, as represented in Fig. 12. Sections 3.1 and 3.2 discuss the structures and details
 219 of the neural networks.

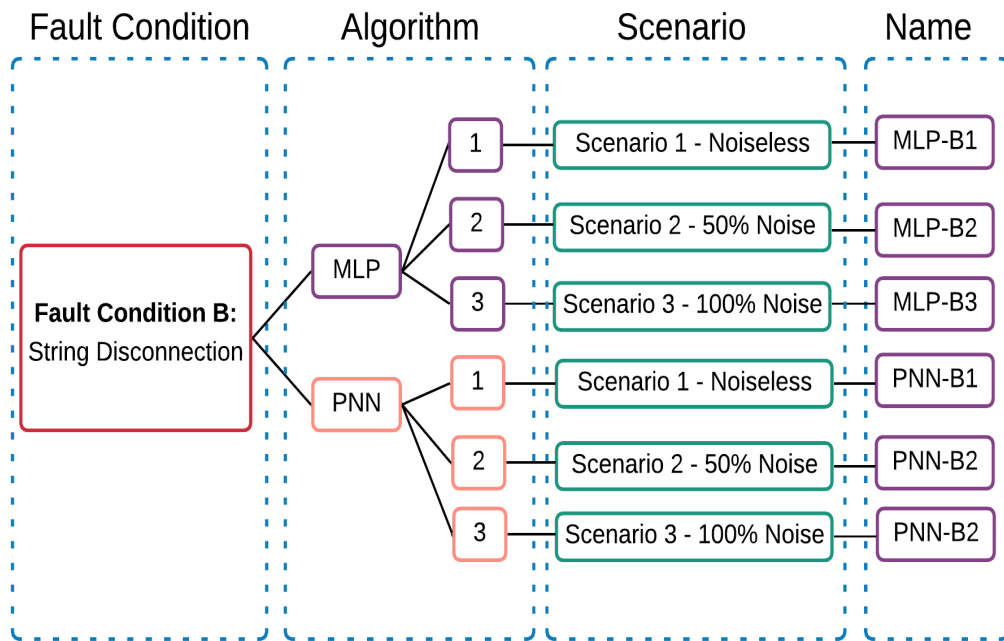


Fig. 12 - Schematic of the studied algorithms and conditions for detecting disconnected strings

220 We also trained three networks for the string disconnection fault situation, one for each studied
 221 scenario, using the same structure in all cases. Fig. 13 shows the neural nets' structure developed using
 222 MATLAB® software, and Table 9 describes its training settings.

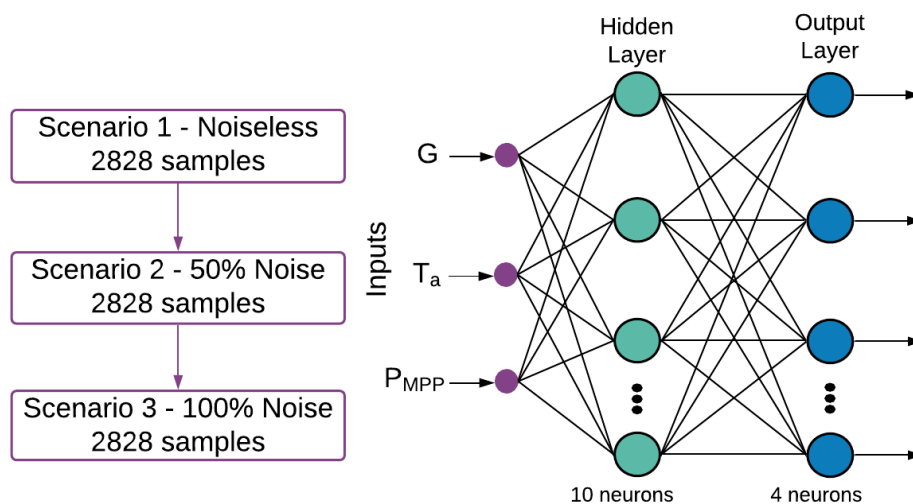


Fig. 13 - MLP network structure for detecting string disconnection

Table 9 - MLP training characteristics for string disconnection detection

MLP Settings	
Input Variables	3 (G, T, P _{MPP})

Output Variables	4
Number of Layers	2
Number of Neurons	(10, 4)
Training Process	supervised
Training Algorithm	Levenberg-Marquardt
Activation Function	(tansingmoid, tansingmoid)
Training	70%
Validation	15%
Test	15%
Type of Divison Samples	random

223 The input variables are equal to those used for the short-circuit fault, and the output vector follows
224 the same logic. For System 2, there are four faulty classes. The first one represents normal operation.
225 Table 10 represents the output vectors for the trained MLPs.

Table 10 - Output vectors for System 2 MLPs

Disconnected Strings	Fault	Output	Class
Normal Operation	F0	1 0 0 0	1
1	F1	0 1 0 0	2
2	F2	0 0 1 0	3
3	F3	0 0 0 1	4

226 Table 11 describes the attributes for the MLPs network training process. Observing the training
227 accuracy results, the MLP-B1 showed an accuracy of 100% on training, while MLP-B2 and B3 showed
228 97.2% and 95.1%, respectively. When we insert the dataset's noise, such as the short-circuit PV
229 modules fault condition, the training accuracy decreases.

Table 11 - MLPs training attributes for detecting disconnected strings

MLP	Epochs	Regression Coefficient	Training Accuracy
B1	48	0.99077	100.0%
B2	28	0.97380	97.2%
B3	36	0.95158	95.1%

230 We also trained three Probabilistic Neural Networks, namely PNN-B1, PNN-B2, and PNN-B3,
231 considering the established scenarios. Fig. 14 shows the neural nets' structure developed using
232 MATLAB® software, and three scenarios were selected as follows:

- 233 • Scenario 1: noiseless samples
- 234 • Scenario 2: 50% of the samples are noisy
- 235 • Scenario 3: 100% of the samples are noisy

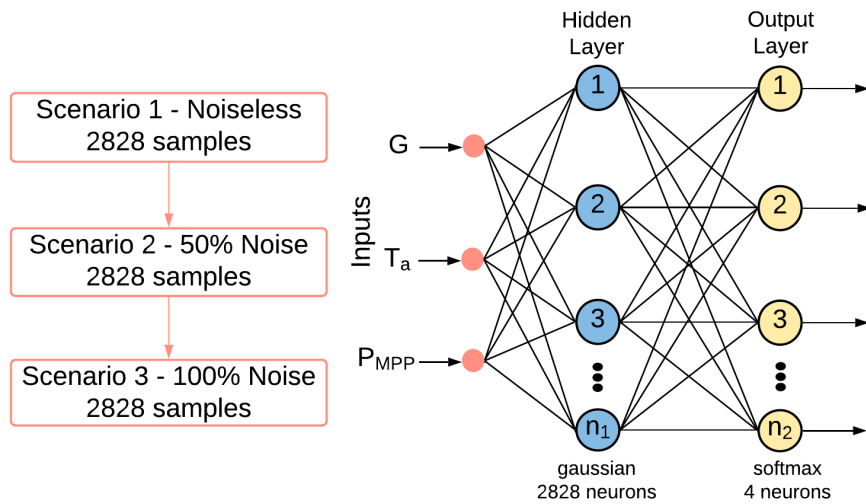


Fig. 14 - PNN network structure for detecting string disconnection

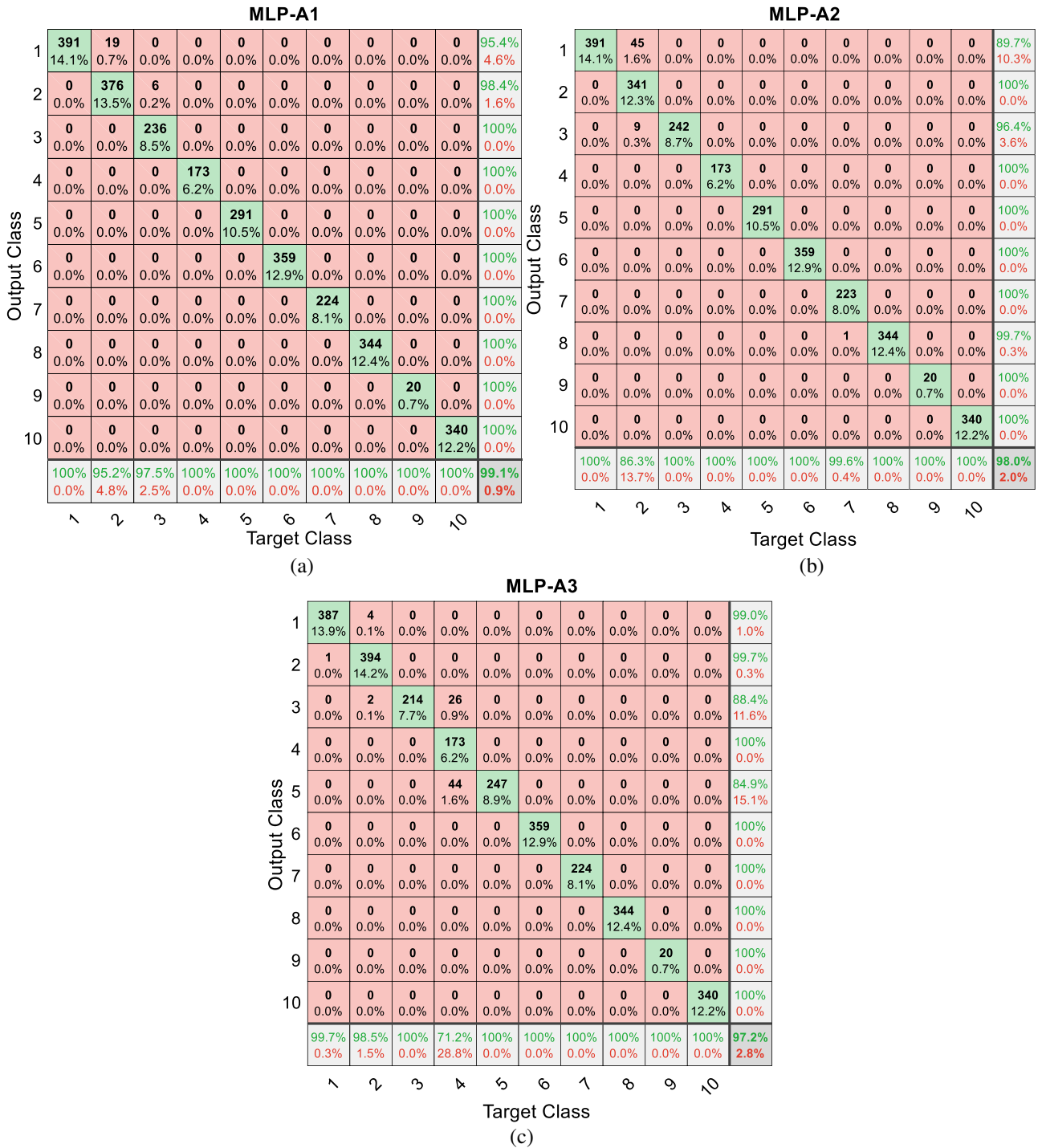
236 The input and output variables are equal to those used for the MLP networks for string
 237 disconnection, following the same output vector described in Table 10. After developing the fault
 238 detection algorithms, it is possible to test the method using experimental results, as discussed in Section
 239 3.

3. Results and Discussion

240 In this section, we will present and discuss the analyzes of the proposed algorithms under field
 241 conditions. Using the experimental results presented in Section 2.1, we tested the developed algorithms
 242 to evaluate their efficiency under real faulty situations. Therefore, Sections 3.1 and 3.2 describe and
 243 discuss the validation of proposed methods for the studied systems.

3.1. Detecting Short-Circuited PV Modules: Methods Validation

244 The extracted results shown in Fig. 3 and Fig. 4 enabled testing the proposed fault detection
 245 methods. We tested the algorithms for short-circuit detection using 2778 experimental samples,
 246 comprising all faulty simulations tackled by the method. Fig. 15 and Fig. 16 show confusion matrices
 247 for the experimental result for the developed neural networks MLP and PNN, respectively.



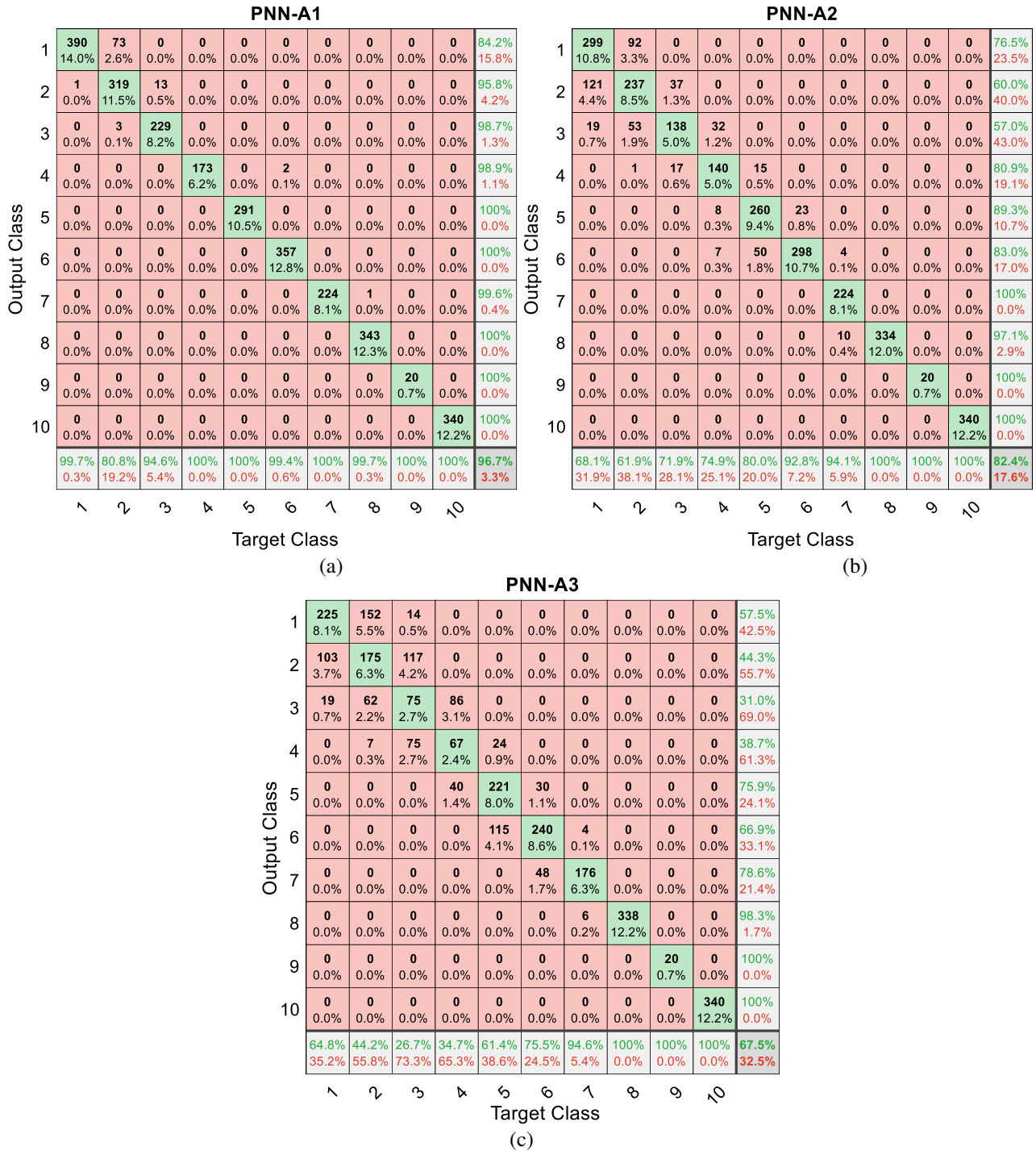


Fig. 16 - System 1 experimental testing confusion matrix (a) PNN-A1, (b) PNN-A2, and (c) PNN-A3

248 To make more precise the results of the experimental tests, we summarized them in Table 12.
 249 Analyzing Table 12, we observe that the MLP algorithm shows a remarkable accuracy of 99.1% for
 250 detecting short-circuited PV modules when trained with a noiseless dataset (MLP-A1). As we insert
 251 the $\pm 15\%$ noise on the MPP data, the accuracy slightly drops, reaching 98% for MLP-A2 and 97.2%
 252 for MLP-A3.

Table 12 – System 1 experimental results on detecting short-circuited PV modules

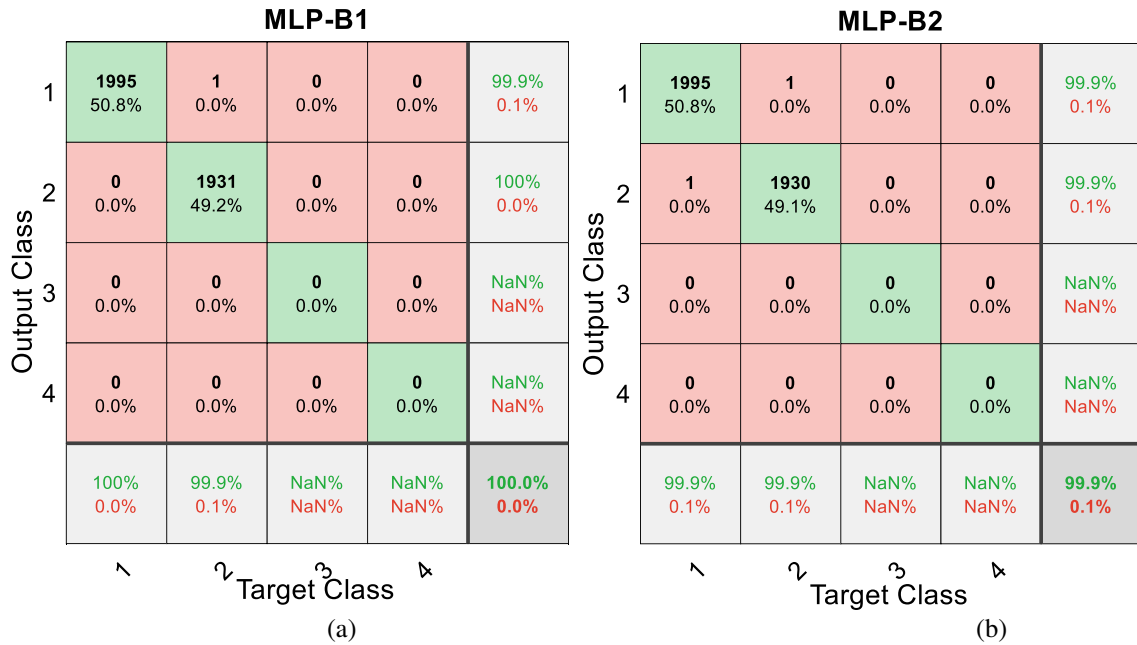
Fault Condition	Algorithm	Scenario	Name	Testing Accuracy
-----------------	-----------	----------	------	------------------

Short-Circuited PV Module	MLP	1 (noiseless)	MLP-A1	99.1%
		2 (50% noise)	MLP-A2	98.0%
		3 (100% noise)	MLP-A3	97.2%
	PNN	1 (noiseless)	PNN-A1	96.7%
		2 (50% noise)	PNN-A2	82.4%
		3 (100% noise)	PNN-A3	67.5%

253 When we compare the MLP algorithm to the PNN, we observe that, in general, the MLP shows
 254 superior accuracy in detecting short-circuited PV modules in all examined scenarios. It is worth
 255 highlighting that the PNN accuracy decays about 29% when trained with the noisy datasets (PNN-A2
 256 and A3). This result reinforces the MLP robustness when the input data is contaminated with random
 257 noises (Lee & Oh, 1994).

3.2. Detecting Disconnected Strings: Methods Validation

258 The extracted results shown in Fig. 6 enabled testing the proposed fault detection methods. For
 259 System 2, we tested the proposed neural networks for string disconnection detection using 3927
 260 experimental samples, comprising normal operation and one string disconnected. The confusion
 261 matrices in Fig. 17 and Fig. 18 show the experimental results for System 2.



MLP-B3

Output Class	1	1995 50.8%	1 0.0%	0 0.0%	0 0.0%	99.9% 0.1%
	2	0 0.0%	1931 49.2%	0 0.0%	0 0.0%	100% 0.0%
	3	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	4	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
		100% 0.0%	99.9% 0.1%	NaN% NaN%	NaN% NaN%	100.0% 0.0%
	1	2	3	4	Target Class	

(c)

Fig. 17 - System 2 experimental testing confusion matrix (a) MLP-B1, (b) MLP-B2, and (c) MLP-B3

PNN-B1						PNN-B2						
Output Class	1	1995 50.8%	1 0.0%	0 0.0%	0 0.0%	99.9% 0.1%	1	1948 49.6%	79 2.0%	0 0.0%	0 0.0%	96.1% 3.9%
	2	21 0.5%	1910 48.6%	0 0.0%	0 0.0%	98.9% 1.1%	2	48 1.2%	1852 47.2%	0 0.0%	0 0.0%	97.5% 2.5%
	3	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%	3	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	4	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%	4	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
		99.0% 1.0%	99.9% 0.1%	NaN% NaN%	NaN% NaN%	99.4% 0.6%		97.6% 2.4%	95.9% 4.1%	NaN% NaN%	NaN% NaN%	96.8% 3.2%
	1	2	3	4	Target Class		1	2	3	4	Target Class	

(a) (b)

PNN-B3

Output Class	1	1942 49.5%	253 6.4%	0 0.0%	0 0.0%	88.5% 11.5%
	2	54 1.4%	1678 42.7%	0 0.0%	0 0.0%	96.9% 3.1%
	3	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
	4	0 0.0%	0 0.0%	0 0.0%	0 0.0%	NaN% NaN%
			97.3% 2.7%	86.9% 13.1%	NaN% NaN%	NaN% NaN%
		Target Class				

(c)

262 Fig. 18 - System 2 experimental testing confusion matrix (a) PNN-B1, (b) PNN-B2, and (c) PNN-B3

263 Table 13 summarizes the results for System 2, making our analyses easier. From Table 13, we
 264 observe that System 2's results follow the same findings for System 1. The MLP algorithms show an
 265 exceptional accuracy of approximately 100% in all examined scenarios, while for the PNN network,
 266 the highest accuracy is 99.4% (PNN-A). The PNN accuracy drops when we insert the noise in the
 267 training dataset, even though, in this case, it decreases less.

Table 13 – System 2 experimental results on detecting disconnected string

Fault Condition	Algorithm	Scenario	Name	Testing Accuracy
String Disconnection	MLP	1 (noiseless)	MLP-B1	100.0%
		2 (50% noise)	MLP-B2	99.9%
		3 (100% noise)	MLP-B3	100.0%
	PNN	1 (noiseless)	PNN-B1	99.4%
		2 (50% noise)	PNN-B2	96.8%
		3 (100% noise)	PNN-B3	92.2%

268 In short, once again, the MLP algorithm showed better accuracy in detecting faulty conditions on
 269 PV systems, as well as it is more robust when considering noisy situations. These results lead us to
 270 conclude that the MLP neural network showed better performance in the analyzed situations, so it is
 271 more suitable for detecting fault occurrence on PV systems.

272 Table 14 indicates the results of the experimental tests performed using the proposed algorithms.
 273 In short, for both tested systems, the MLP neural network showed superior accuracy than PNN.
 274 Furthermore, the MLP algorithms showed superior accuracy in all examined situations than the PNN
 275 and were more robust to noisy training datasets. Thus, it makes the algorithm not only more accurate
 276 but also more reliable.

Table 14 – Experimental results for the proposed fault detection method

Fault Condition	Algorithm	Scenario	Name	Testing Accuracy
Short-Circuited PV Module	MLP	1 (noiseless)	MLP-A1	99.1%
		2 (50% noise)	MLP-A2	98.0%
		3 (100% noise)	MLP-A3	97.2%
	PNN	1 (noiseless)	PNN-A1	96.7%
		2 (50% noise)	PNN-A2	82.4%
		3 (100% noise)	PNN-A3	67.5%
String Disconnection	MLP	1 (noiseless)	MLP-B1	100.0%
		2 (50% noise)	MLP-B2	99.9%
		3 (100% noise)	MLP-B3	100.0%
	PNN	1 (noiseless)	PNN-B1	99.4%
		2 (50% noise)	PNN-B2	96.8%
		3 (100% noise)	PNN-B3	92.2%

For detecting short-circuited PV modules on System 1, the trained MLP showed the highest accuracy of 99.1% for the noiseless condition (MLP-A1) and decreased to 98% (MLP-A2) and 97.2% (MLP-A3) when we considered noisy scenarios 2 and 3. In System 2, the MLP detected disconnected strings, presenting a remarkable accuracy of approximately 100% in all examined situations.

3.3. Comparative Study

277 To assess the proposed research findings over previously published studies, we developed
278 Table 15, presenting the type of detected fault, algorithm, and method's accuracy.

Table 15 – Comparison with previously published research

Reference	Fault	Algorithm	Accuracy	Experimentally tested
(Chao, Chen, Wang, & Wu, 2010)	<ul style="list-style-type: none"> Faulty Modules 	MLP	93.33%	No
(Akram & Lotfifard, 2015)	<ul style="list-style-type: none"> Open circuit module Short-circuited modules 	PNN	96.50%	No
(Chine et al., 2016)	<ul style="list-style-type: none"> Partial shading Bypass diode Short-circuited modules 	MLP	90.30%	Yes
		RBF	68.40%	
(Garoudja et al., 2017)	<ul style="list-style-type: none"> Short-circuited PV modules Disconnected strings 	MLP	90.30%	Yes
		PNN	100.00%	
(Madeti & Singh, 2018)	<ul style="list-style-type: none"> Open circuit module Line to line fault Shading Bypass diode 	kNN	98.70%	No

(Dhimish et al., 2018)	<ul style="list-style-type: none"> • Partial shading • Faulty PV module 	RBF	92.10%	Yes
(Hussain, Dhimish, Titarenko, & Mather, 2020)	<ul style="list-style-type: none"> • Short-circuited PV module • String disconnection 	MLP	97.00%	Yes

279 To make a reasonable comparison, we mentioned those researches that applied an ANN
280 algorithm and detected faults equal or comparable to those approached in this study. Unfortunately,
281 none of the referenced studies in Table 15 considers noisy data on its training, so we are considering
282 the results with noiseless datasets for this analysis.

283 As discussed in Sections 3.1 and 3.2, the MLP's algorithms showed the best accuracy in the
284 context of this research. The results indicated 99.1% (MLP-A1) correctness on detecting short-
285 circuited PV modules and 100% (MLP-B1) detecting disconnected strings. Compared to the research
286 underlined in Table 11, the obtained results indicated the highest accuracy.

287 It is well established that the performance of neural networks depends on the quality of the
288 training data (Kordos & Rusiecki, 2016). However, this research demonstrated that even when using
289 flawed training datasets, the MLP network comes out with excellent accuracy of 97% (MLP-A3),
290 equivalent to those presented in Table 15.

291 Particularly compared to Garoudja *et al.* (2017), which also developed MLP and PNN networks
292 to detect faults on PV modules, we can highlight that the method proposed in this study identifies how
293 many modules or strings are on faulty conditions. Besides, it requires fewer input variables.

294

4. Final Remarks

295 This paper compares MLP and PNN neural networks for detecting faults occurring on a PV
296 system. We trained both algorithms using simulated datasets and considered three different scenarios.
297 For the first situation, Scenario 1, we used the raw data extracted by simulation. In the other two
298 situations, named Scenario 1 and 2, we inserted a $\pm 15\%$ noise on the P_{MPP} data. This noise represents
299 the uncertainties associated with the MPPT device.

300 The analyzed conditions make the method suitable to any PV plant, considering it does not
301 require data from pre-existing systems. It basically needs to retrain the ANN. The input variables are
302 irradiance, ambient temperature, and power at the maximum power point. The ANNs output is a vector
303 indicating which fault is occurring on the PV system. The faults identified by the proposed methods
304 are short-circuited PV modules and disconnected strings.

305 We tested the MLP and PNN neural networks using experimental data from two PV systems
306 installed on the Huddersfield University campus. The first one, named here as System 1, comprises a

307 2.2 kW_p PV system. The second system, named System 2, is a 4.16 kW_p PV system. The results
308 indicated superior accuracy of the MLP algorithm in all examined conditions, especially when
309 considering the noisy datasets. These findings reinforced the robustness of MLP neural nets for pattern
310 recognition, even when the training data is flawed. Furthermore, the noise insertion was not studied
311 before in the current state-of-the-art, thus launching an essential prospect for future researches.

312 The main limitation of the proposed method involves retraining the ANN to be implemented
313 on any PV system. Besides, it requires specific training data for each system, according to the
314 characteristics of the plant. So, there is a need for developing a flexible model that could be employed
315 in any PV system with minor modifications.

316 These findings allowed us to conclude that the MLP neural network is more suitable than PNNs
317 for PV system fault detection, even when the data is contaminated with random noise.

References

- 318 Akram, M. N., & Lotfifard, S. (2015). Modeling and Health Monitoring of DC Side of Photovoltaic
319 Array. *IEEE Transactions on Sustainable Energy*, 1–9.
- 320 Al-Atrash, H., Batarseh, I., & Rustom, K. (2010). Effect of measurement noise and bias on hill-
321 climbing MPPT algorithms. *IEEE Transactions on Aerospace and Electronic Systems*, 46(2),
322 745–760. <https://doi.org/10.1109/TAES.2010.5461654>
- 323 Chao, K. H., Chen, C. T., Wang, M. H., & Wu, C. F. (2010). A novel fault diagnosis method based-
324 on modified neural networks for photovoltaic systems. *ICSI: International Conference in Swarm
325 Intelligence, Part II*, 531–539. https://doi.org/10.1007/978-3-642-13498-2_69
- 326 Chine, W., Mellit, A., Lughy, V., Malek, A., Sulligoi, G., & Massi Pavan, A. (2016). A Novel Fault
327 Diagnosis Technique for Photovoltaic Systems Based on Artificial Neural Networks. *Renewable
328 Energy*, 90, 501–512. <https://doi.org/10.1016/j.renene.2016.01.036>
- 329 Dhimish, M., Holmes, V., Mehrdadi, B., & Dales, M. (2018). Comparing Mamdani Sugeno fuzzy logic
330 and RBF ANN network for PV fault detection. *Renewable Energy*, 117, 257–274.
331 <https://doi.org/10.1016/j.renene.2017.10.066>
- 332 Dhimish, M., Hu, Y., Schofield, N., & Vieira, R. G. (2020). Mitigating Potential-Induced Degradation
333 (PID) Using SiO₂ ARC Layer. *Energies*, 13(19), 1–12.
- 334 Garoudja, E., Chouder, A., Kara, K., & Silvestre, S. (2017). An Enhanced Machine Learning Based
335 Approach for Failures Detection and Diagnosis of PV Systems. *Energy Conversion and
336 Management*, 151(September), 496–513. <https://doi.org/10.1016/j.enconman.2017.09.019>
- 337 Ghaffarzadeh, N., & Azadian, A. (2019). A Comprehensive Review and Performance Evaluation in
338 Solar (PV) Systems Fault Classification and Fault Detection Techniques. *Journal of Solar Energy*

339 *Research*, 4(4), 252–272.

340 Guerra, M. I. S., Ara, M. U. De, Dhimish, M., & Vieira, G. (2021). Assessing Maximum Power Point
341 Tracking Intelligent Techniques on a PV System with a Buck – Boost Converter. *Energies*.

342 Hussain, M., Dhimish, M., Titarenko, S., & Mather, P. (2020). Artificial Neural Network Based
343 Photovoltaic Fault Detection Algorithm Integrating Two Bi-Directional Input Parameters.
344 *Renewable Energy*, 155, 1272–1292. <https://doi.org/10.1016/j.renene.2020.04.023>

345 IEA. (2020). *Snapshot of Global Photovoltaic Markets*. Retrieved from [http://www.iaepvps.org/fileadmin/dam/public/report/technical/PVPS_report_-_A_Snapshot_of_Global_PV_-_](http://www.iaepvps.org/fileadmin/dam/public/report/technical/PVPS_report_-_A_Snapshot_of_Global_PV_-_1992-2014.pdf)
346 [_1992-2014.pdf](http://www.iaepvps.org/fileadmin/dam/public/report/technical/PVPS_report_-_A_Snapshot_of_Global_PV_-_1992-2014.pdf)

347

348 Jiang, L. L., & Maskell, D. L. (2015). Automatic Fault Detection and Diagnosis for Photovoltaic
349 Systems Using Combined Artificial Neural Network and Analytical Based Methods. *Proceedings*
350 *of the International Joint Conference on Neural Networks, 2015-Sept.*
351 <https://doi.org/10.1109/IJCNN.2015.7280498>

352 Kordos, M., & Rusiecki, A. (2016). Reducing noise impact on MLP training: Techniques and
353 algorithms to provide noise-robustness in MLP network training. *Soft Computing*, 20(1), 49–65.
354 <https://doi.org/10.1007/s00500-015-1690-9>

355 Lee, Y., & Oh, S. H. (1994). Input noise immunity of multilayer perceptrons. *ETRI Journal*, 16(1),
356 35–43. <https://doi.org/10.4218/etrij.94.0194.0013>

357 Li, Z., Wang, Y., Zhou, D., & Wu, C. (2012). An Intelligent Method for Fault Diagnosis in
358 Photovoltaic Array. *International Conference on Electrical and Information Technologies*
359 (*ICEIT*), 10–16. Rabat, Morocco: IEEE, 15-18 November.

360 Liu, Y., Zhu, X., & Yang, J. (2017). Fault Diagnosis of PV Array Based on Optimised BP Neural
361 Network By Improved Adaptive Genetic Algorithm. *The Journal of Engineering*, 2017(13),
362 1427–1431. <https://doi.org/10.1049/joe.2017.0567>

363 Madeti, S. R., & Singh, S. N. (2017). A Comprehensive Study on Different Types of Faults and
364 Detection Techniques for Solar Photovoltaic System. *Solar Energy*, 158(August), 161–185.
365 <https://doi.org/10.1016/j.solener.2017.08.069>

366 Madeti, S. R., & Singh, S. N. (2018). Modeling of PV System Based on Experimental Data for Fault
367 Detection Using kNN Method. *Solar Energy*, 173(March), 139–151.
368 <https://doi.org/10.1016/j.solener.2018.07.038>

369 Pillai, D. S., Blaabjerg, F., & Rajasekar, N. (2019). A Comparative Evaluation of Advanced Fault
370 Detection Approaches for PV Systems. *IEEE Journal of Photovoltaics*, 9(2), 513–527.
371 <https://doi.org/10.1109/JPHOTOV.2019.2892189>

372 Siddique, N., & Adeli, H. (2013). *Computational Intelligence Computational Intelligence Synergies*

- 373 *of Fuzzy Logic, Neural Networks and Evolutionary Computing* (1st ed.). New Delhi, India: John
374 Wiley & Sons.
- 375 Syafaruddin, Karatepe, E., & Hiyama, T. (2011). Controlling of artificial neural network for fault
376 diagnosis of photovoltaic array. *16th International Conference on Intelligent System Applications*
377 *to Power Systems*, 1–6. <https://doi.org/10.1109/ISAP.2011.6082219>
- 378 Vieira, R. G., de Araújo, F. M. U., Dhimish, M., & Guerra, M. I. S. (2020). A Comprehensive Review
379 on Bypass Diode Application on Photovoltaic Modules. *Energies*, *13*(10), 2472.
380 <https://doi.org/10.3390/en13102472>
- 381 Vieira, R. G., Dhimish, M., de Araújo, F. M. U., & Guerra, M. I. S. (2020). PV Module Fault Detection
382 Using Combined Artificial Neural Network and Sugeno Fuzzy Logic. *Electronics (Switzerland)*,
383 *9*(12), 1–21. <https://doi.org/10.3390/electronics9122150>