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

Romênia Gurgel Vieira[1,*], Mahmoud Dhimish[2], Fábio Meneghetti Ugulino de Araújo[3], Maria Izabel da Silva Guerra[1]

[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)

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

Keywords*: Solar Energy; Photovoltaic Modules; String Disconnection; Short-circuit; Fault 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

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

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

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

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

1.1. Literature Review

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

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

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

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

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

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

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

1.2. Related Studies

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

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

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

2. Research Methodology

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

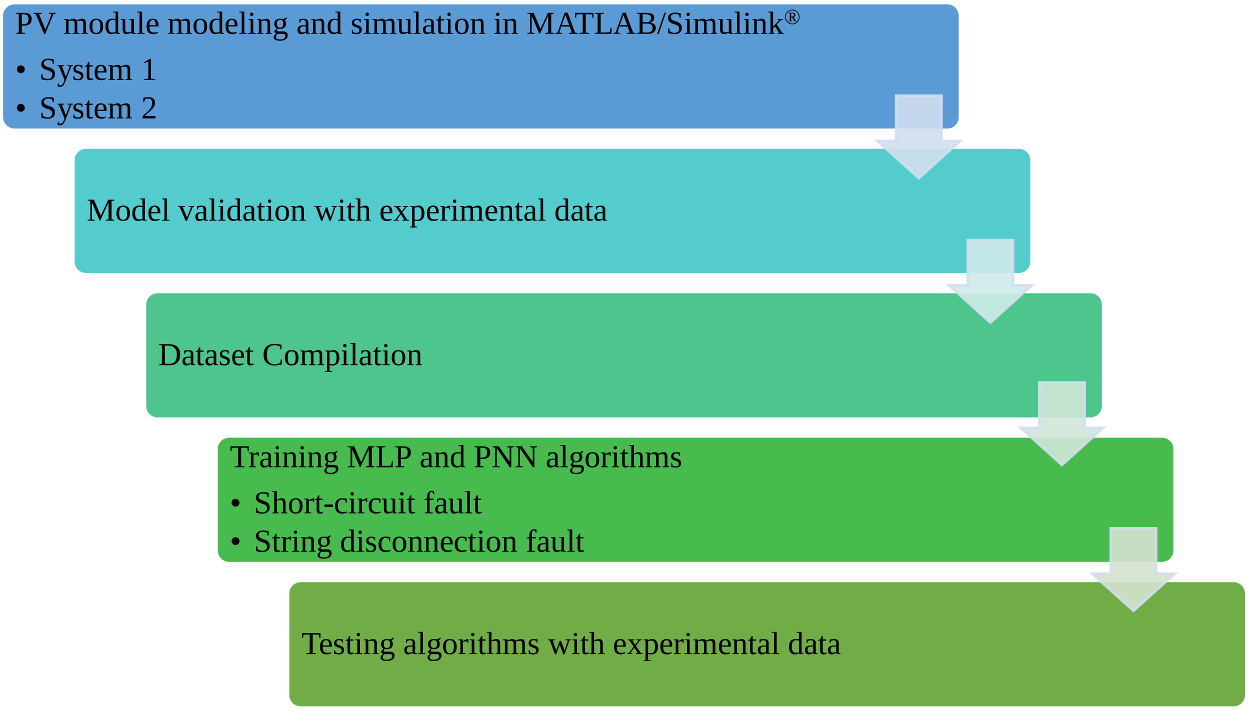


Fig. 1 - Research methodology workflow

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

2.1. System 1 and 2: Description and Model Validation

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

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

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

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

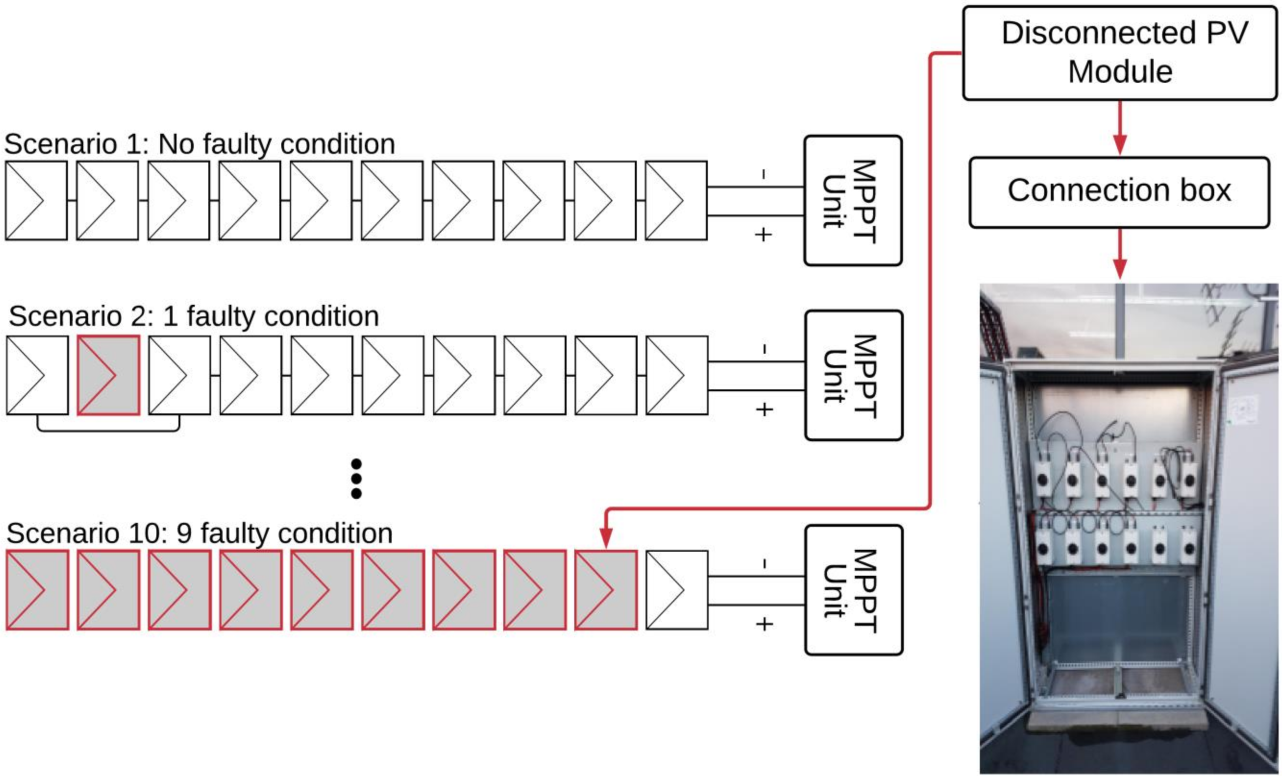


Fig. 2 - System 1 experimental setup

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

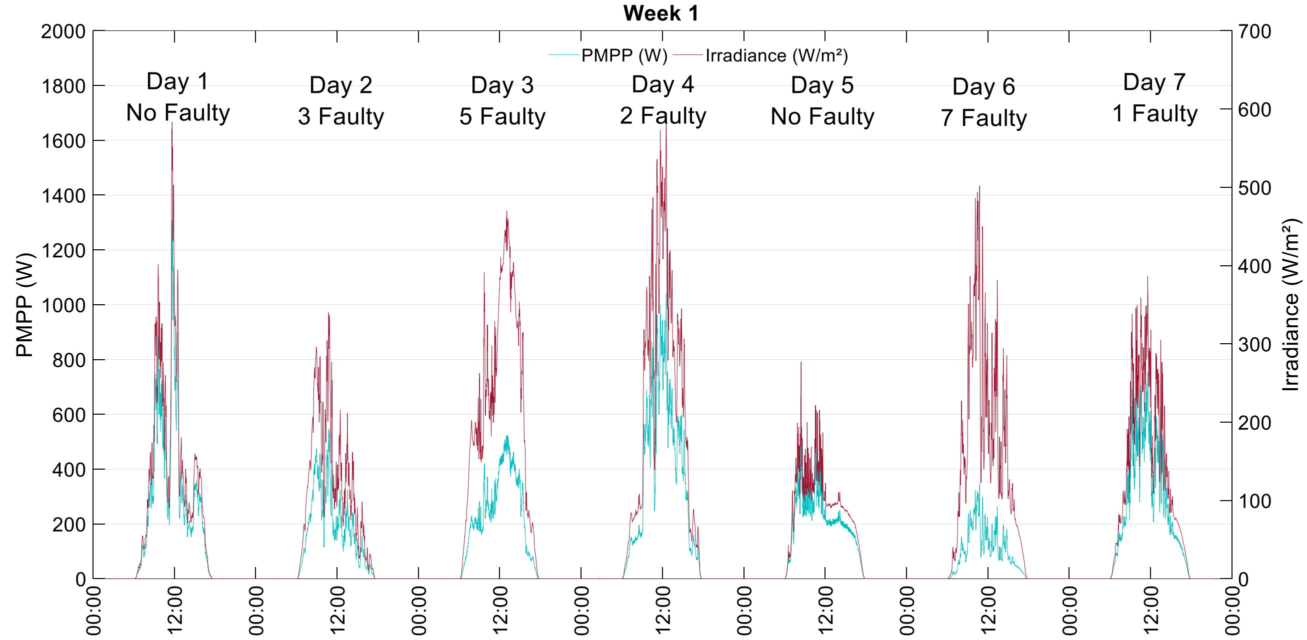


Fig. 3 - Week 1 experimental results for System 1

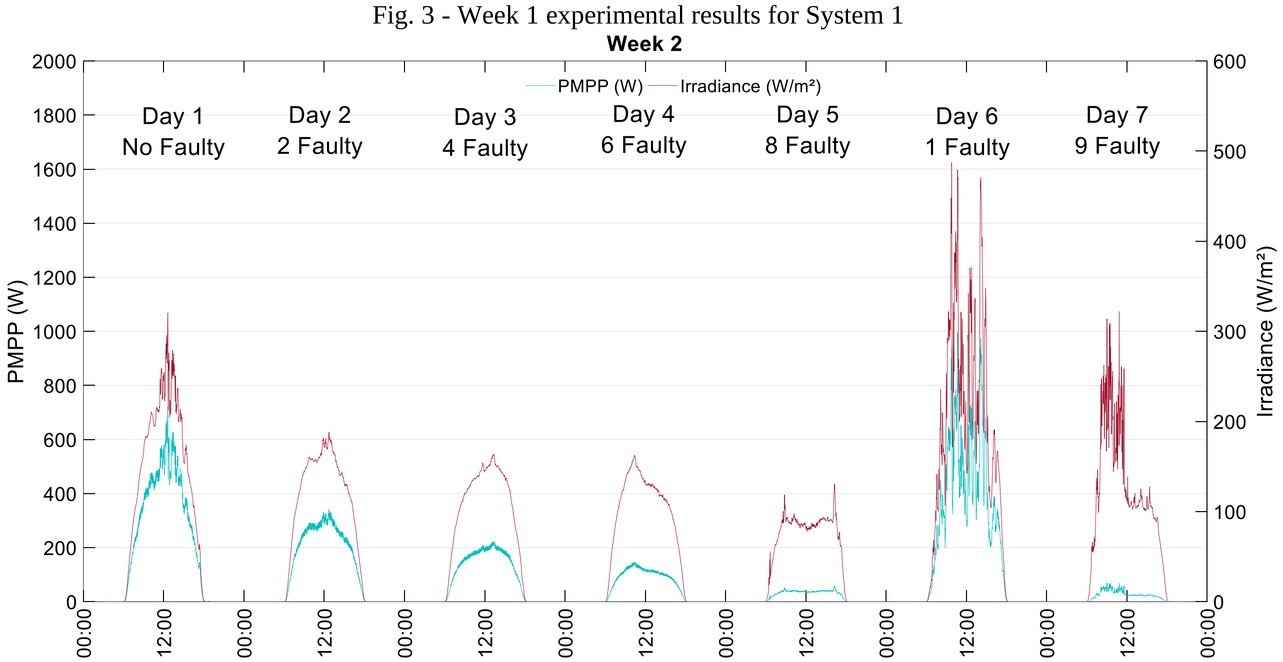


Fig. 4 - Week 2 experimental results for System 1

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

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

Table 3 – System 1 modeling validation

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

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

The second studied PV system is a 4.16 kWp power plant also installed at the Huddersfield University campus. It comprises 32 PV modules arranged into four strings, with eight modules each. The module model is the KC130GHT-2 from Kyocera manufacturer, with a nominal power of 130 W, 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			
$\mathbf{R_{sh}}$	119.232 Ω	$\mathbf{R_s}$	0.16 Ω

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

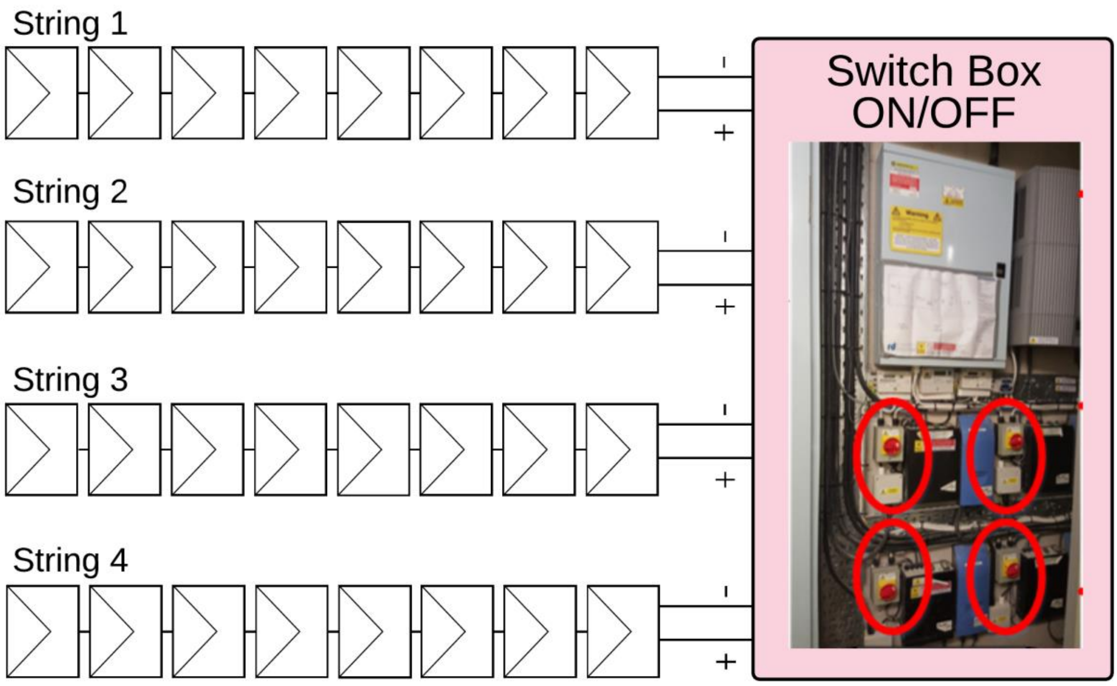


Fig. 5 - System 2 experimental setup

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

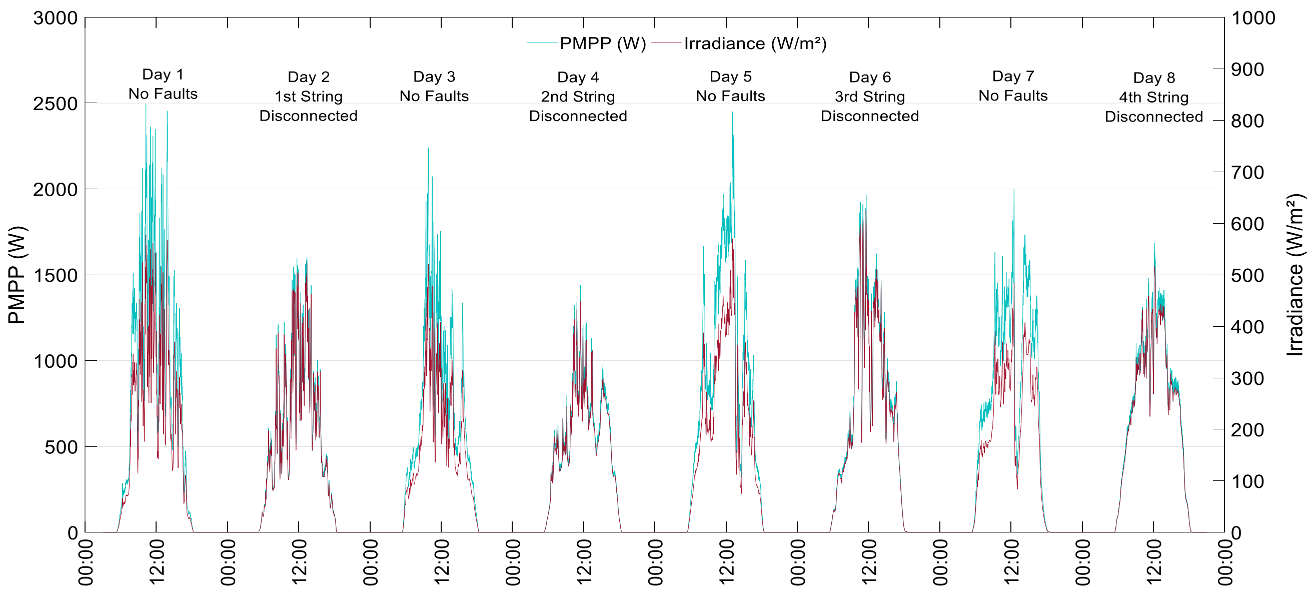


Fig. 6 - Experimental results for System 2

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

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

Table 5 - System 2 modeling validation

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

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

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

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

We developed the algorithms for the short-circuited PV module faulty condition considering System 1. The obtained dataset comprises 7070 samples, 707 for each faulty condition. We settled three scenarios for evaluating the algorithms: Scenario 1, Scenario 2, and Scenario 3. The first case, Scenario 1, corresponds to the raw data extracted by simulation. For the others examined conditions, we inserted a noise of ±15% in the P_{MPP} input variable. Thus, Scenario1 is a noiseless condition, while Scenario 2 contains a noise of ±15% on 50% of the MPP data, and Scenario 3 contains the noise in 100% of the MPP data.

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

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

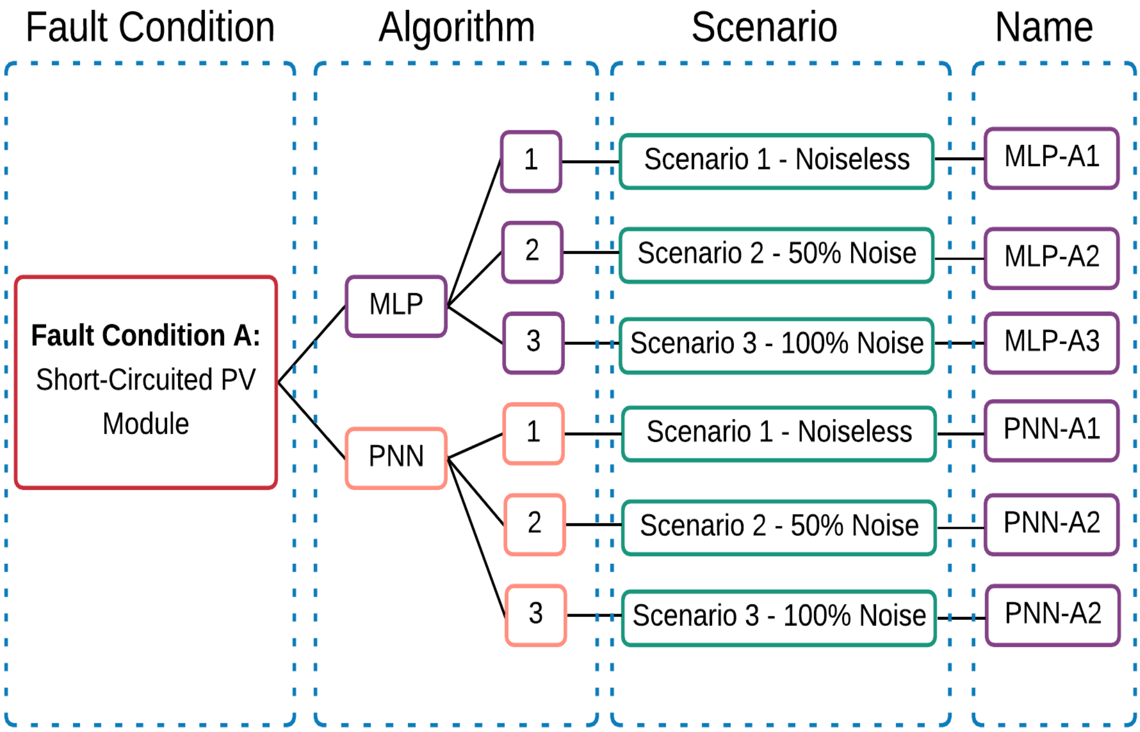


Fig. 7 - Schematic of the studied algorithms and conditions for detecting short-circuited PV modules

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

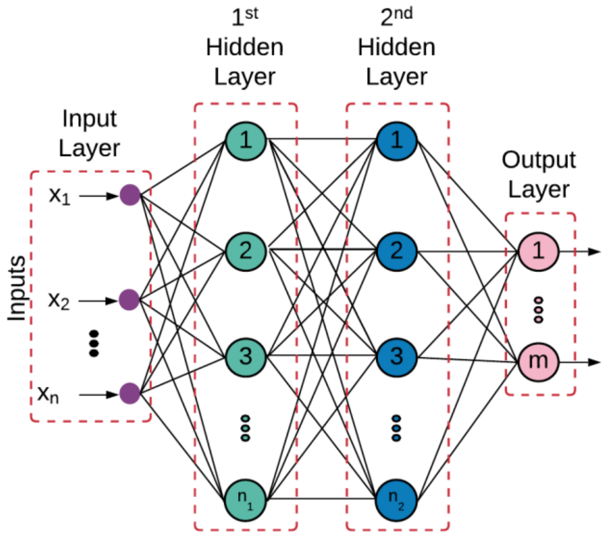


Fig. 8 - MLP basic structure

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

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

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

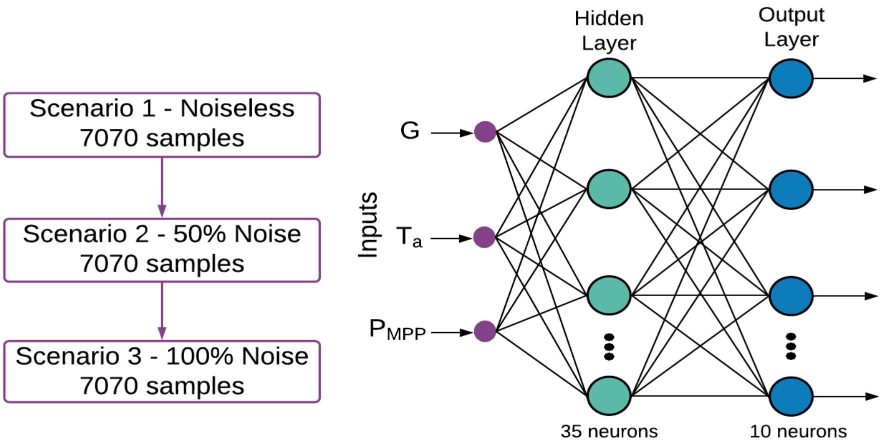


Fig. 9 - MLP network structure for detecting short-circuit PV modules

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

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

The input variables are irradiance (G), ambient temperature (T_a), and the maximum power point (P_{MPP}). The output is a vector equals zero, except for one element equals 1. This element represents the faulty condition identified. For System 1, there are ten faulty classes. The first one represents 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

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

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

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

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

Where i represents the class number, and j the pattern number, x_{ij} is the j^{th} training vector from i, 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 smoothing factor and $f_i(x)$ is the sum of multivariate Gaussian distribution centered at each of the training samples.

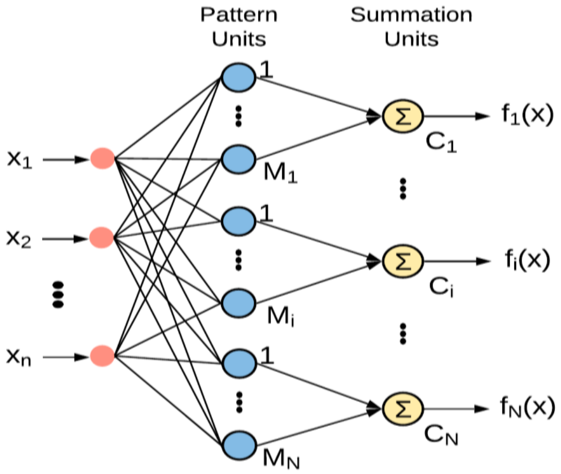


Fig. 10 - PNN basic structure

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

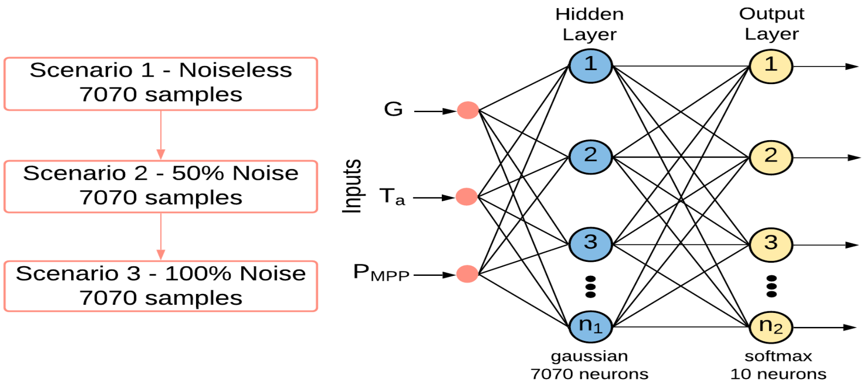


Fig. 11 - PNN network structure for detecting short-circuit PV modules

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

2.3. Detecting Disconnected Strings: MLP and PNN

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

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

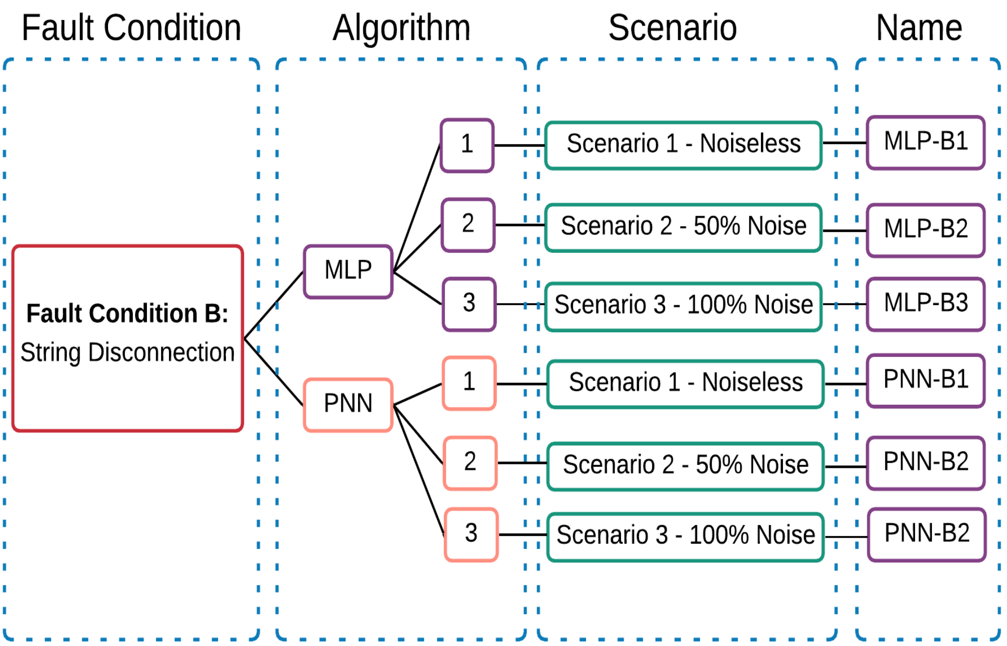


Fig. 12 - Schematic of the studied algorithms and conditions for detecting disconnected strings

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

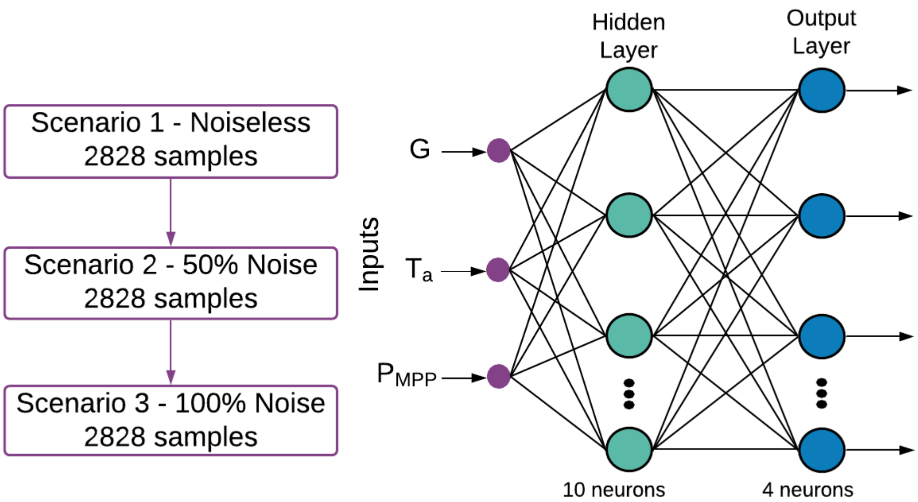


Fig. 13 - MLP network structure for detecting string disconnection

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

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

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

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

- Scenario 1: noiseless samples
- Scenario 2: 50% of the samples are noisy
- Scenario 3: 100% of the samples are noisy

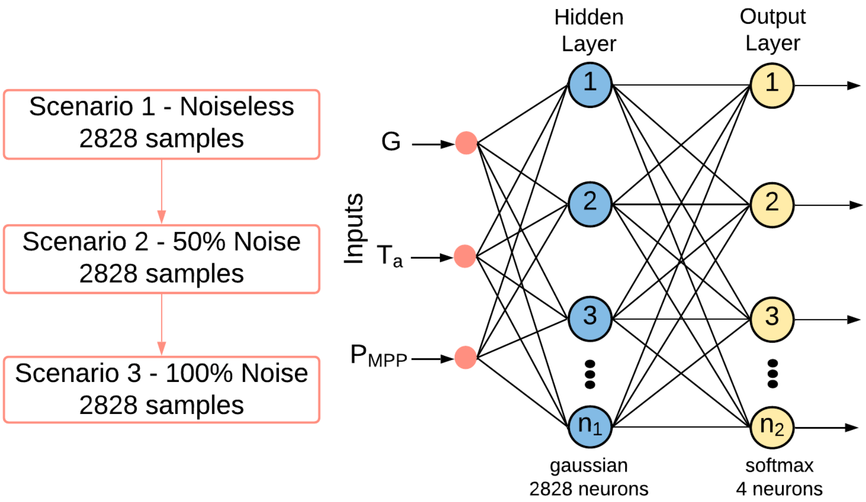


Fig. 14 - PNN network structure for detecting string disconnection

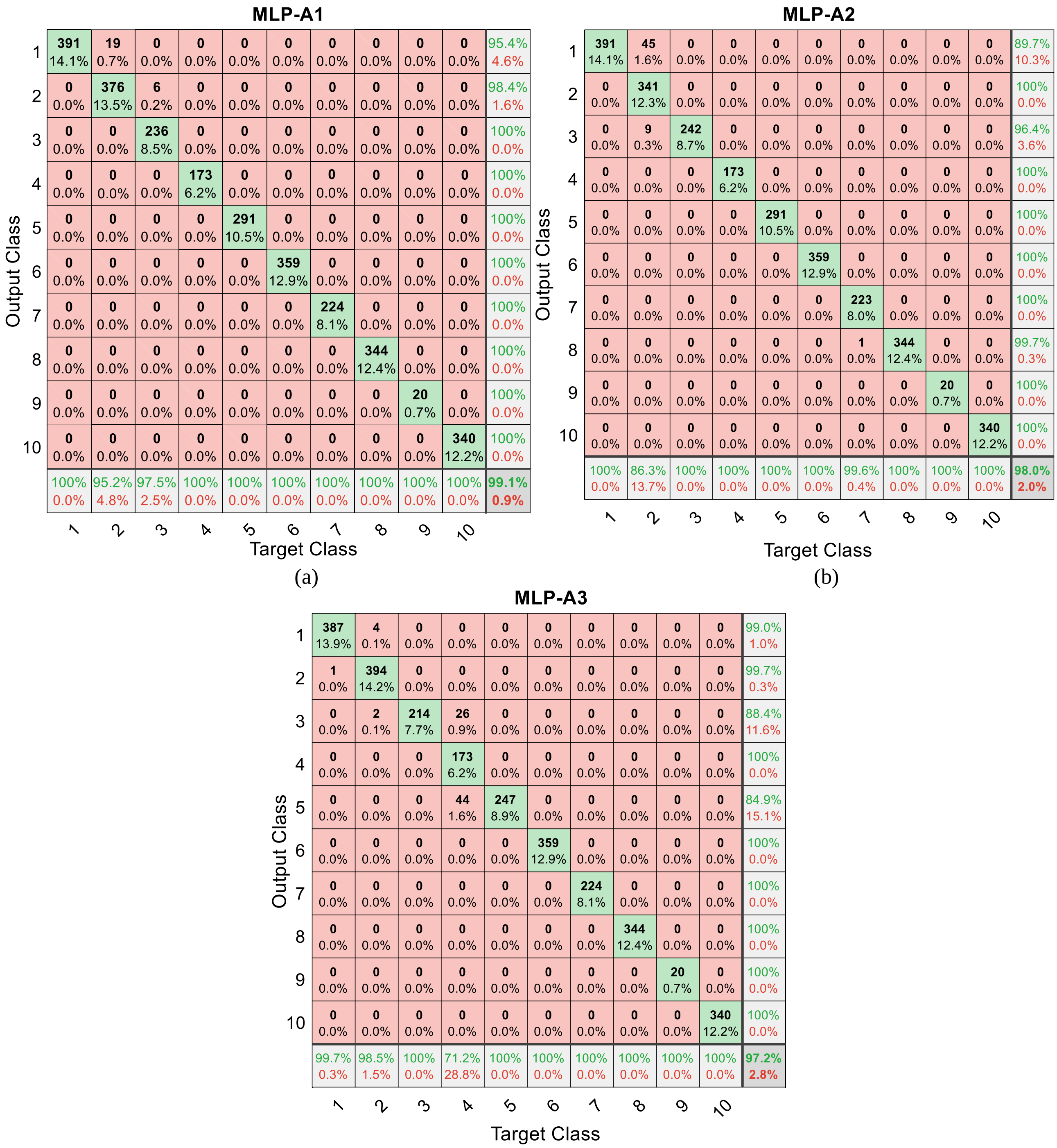
The input and output variables are equal to those used for the MLP networks for string disconnection, following the same output vector described in Table 10. After developing the fault detection algorithms, it is possible to test the method using experimental results, as discussed in Section 3.

3. Results and Discussion

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

3.1. Detecting Short-Circuited PV Modules: Methods Validation

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



(c)

Fig. 15 - System 1 experimental testing confusion matrix (a) MLP-A1, (b) MLP-A2, and (c) MLP-A3

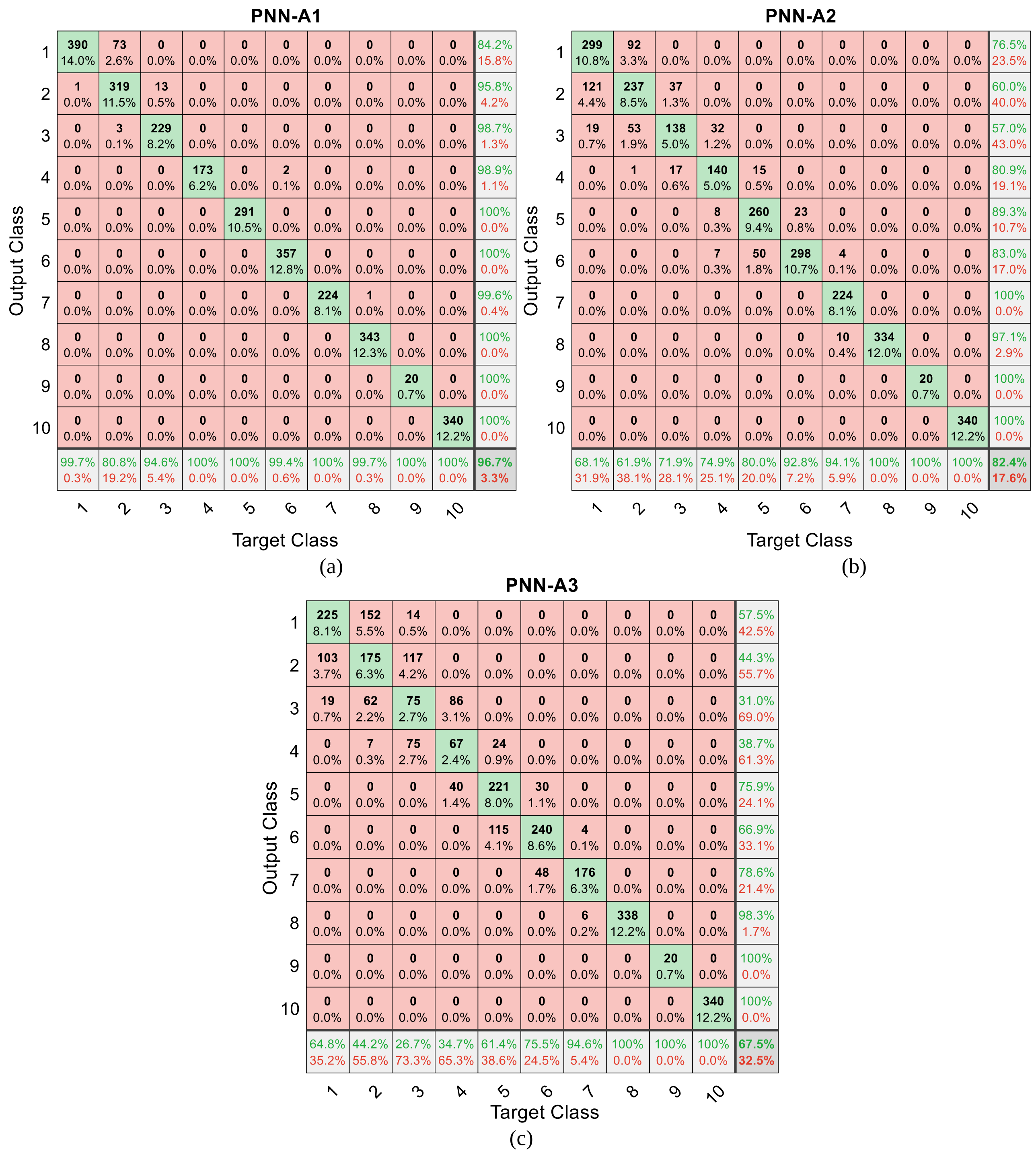


Fig. 16 - System 1 experimental testing confusion matrix (a) PNN-A1, (b) PNN-A2, and (c) PNN-A3

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

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

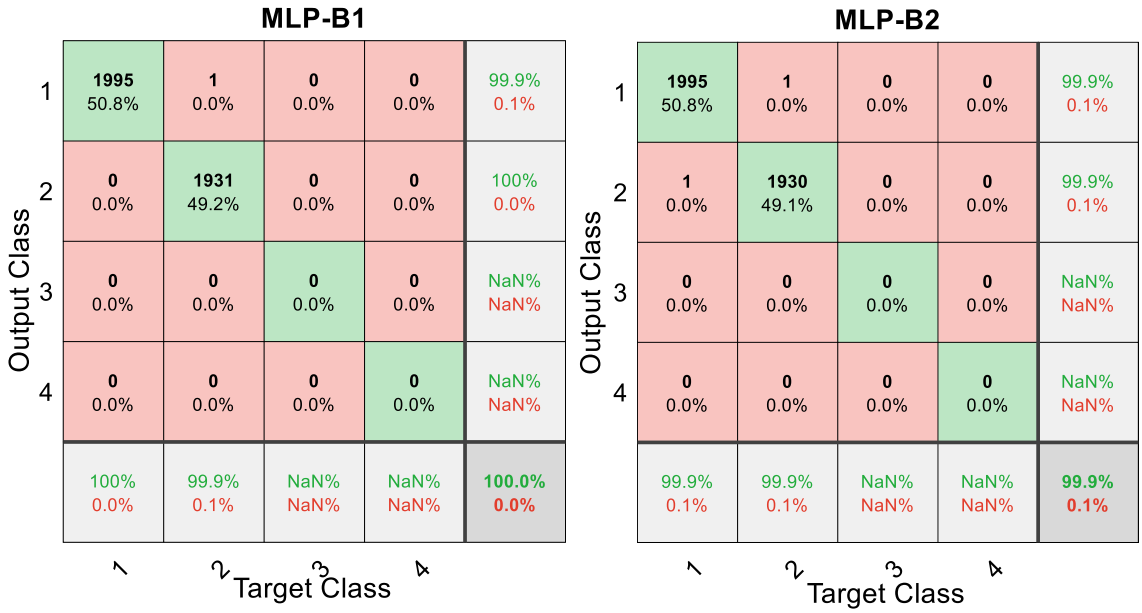
Fault Condition	Algorithm	Scenario	Name	Testing Accuracy

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%

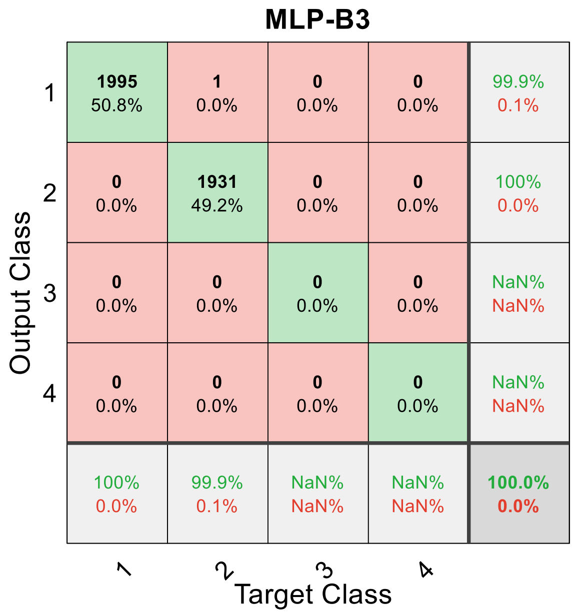
When we compare the MLP algorithm to the PNN, we observe that, in general, the MLP shows superior accuracy in detecting short-circuited PV modules in all examined scenarios. It is worth highlighting that the PNN accuracy decays about 29% when trained with the noisy datasets (PNN-A2 and A3). This result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

3.2. Detecting Disconnected Strings: Methods Validation

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

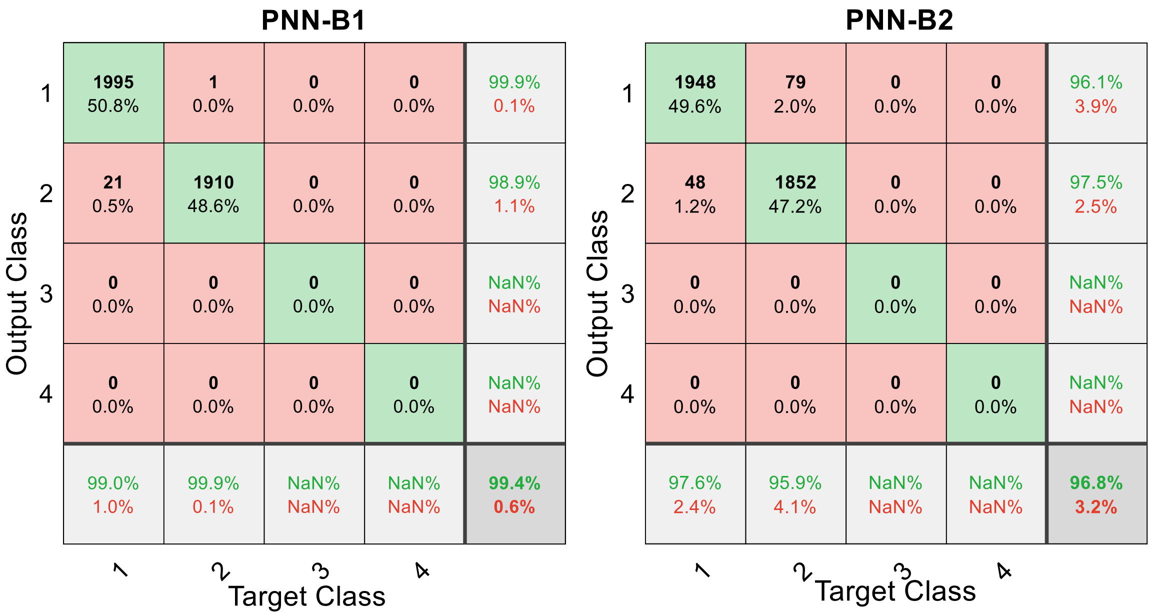


(a) (b)

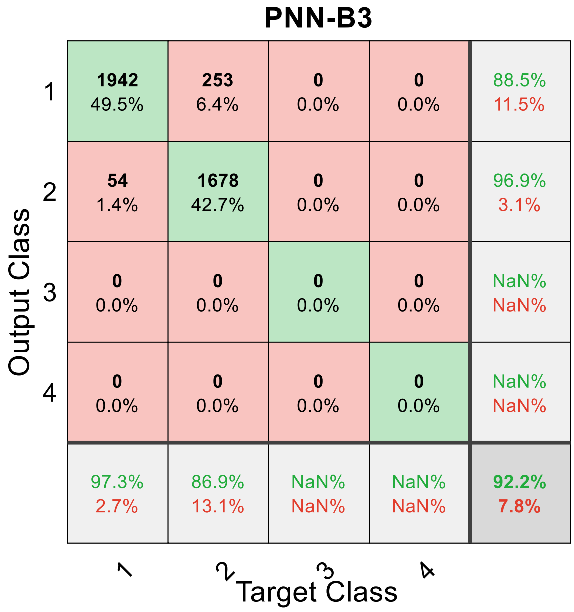


(c)

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



(a) (b)



(c)

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

Table 13 summarizes the results for System 2, making our analyses easier. From Table 13, we observe that System 2's results follow the same findings for System 1. The MLP algorithms show an exceptional accuracy of approximately 100% in all examined scenarios, while for the PNN network, the highest accuracy is 99.4% (PNN-A). The PNN accuracy drops when we insert the noise in the 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%

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

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

To assess the proposed research findings over previously published studies, we developed 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)	• Faulty Modules	MLP	93.33%	No
(Akram & Lotfifard, 2015)	• Open circuit module • Short-circuited modules	PNN	96.50%	No
(Chine et al., 2016)	• Partial shading • Bypass diode • Short-circuited modules	MLP RBF	90.30% 68.40%	Yes
(Garoudja et al., 2017)	• Short-circuited PV modules • Disconnected strings	MLP PNN	90.30% 100.00%	Yes
(Madeti & Singh, 2018)	• Open circuit module • Line to line fault • Shading • Bypass diode	*k*NN	98.70%	No

(Dhimish et al., 2018)	• Partial shading • Faulty PV module	RBF	92.10%	Yes
(Hussain, Dhimish, Titarenko, & Mather, 2020)	• Short-circuited PV module • String disconnection	MLP	97.00%	Yes

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

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

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

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

4. Final Remarks

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

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

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

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

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

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

References

Akram, M. N., & Lotfifard, S. (2015). Modeling and Health Monitoring of DC Side of Photovoltaic Array. *IEEE Transactions on Sustainable Energy*, 1–9.

Al-Atrash, H., Batarseh, I., & Rustom, K. (2010). Effect of measurement noise and bias on hill-climbing MPPT algorithms. *IEEE Transactions on Aerospace and Electronic Systems*, *46*(2), 745–760. https://doi.org/10.1109/TAES.2010.5461654

Chao, K. H., Chen, C. T., Wang, M. H., & Wu, C. F. (2010). A novel fault diagnosis method based-on modified neural networks for photovoltaic systems. *ICSI: International Conference in Swarm Intelligence*, *Part II*, 531–539. https://doi.org/10.1007/978-3-642-13498-2_69

Chine, W., Mellit, A., Lughi, V., Malek, A., Sulligoi, G., & Massi Pavan, A. (2016). A Novel Fault Diagnosis Technique for Photovoltaic Systems Based on Artificial Neural Networks. *Renewable Energy*, *90*, 501–512. https://doi.org/10.1016/j.renene.2016.01.036

Dhimish, M., Holmes, V., Mehrdadi, B., & Dales, M. (2018). Comparing Mamdani Sugeno fuzzy logic and RBF ANN network for PV fault detection. *Renewable Energy*, *117*, 257–274. https://doi.org/10.1016/j.renene.2017.10.066

Dhimish, M., Hu, Y., Schofield, N., & Vieira, R. G. (2020). Mitigating Potential-Induced Degradation (PID) Using SiO 2 ARC Layer. *Energies*, *13*(19), 1–12.

Garoudja, E., Chouder, A., Kara, K., & Silvestre, S. (2017). An Enhanced Machine Learning Based Approach for Failures Detection and Diagnosis of PV Systems. *Energy Conversion and Management*, *151*(September), 496–513. https://doi.org/10.1016/j.enconman.2017.09.019

Ghaffarzadeh, N., & Azadian, A. (2019). A Comprehensive Review and Performance Evaluation in Solar (PV) Systems Fault Classification and Fault Detection Techniques. *Journal of Solar Energy*

Research, *4*(4), 252–272.

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

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

IEA. (2020). *Snapshot of Global Photovoltaic Markets*. Retrieved from http://www.iea-pvps.org/fileadmin/dam/public/report/technical/PVPS_report_-_A_Snapshot_of_Global_PV_-_1992-2014.pdf

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

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

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

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

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

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

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

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

Siddique, N., & Adeli, H. (2013). *Computational Intelligence Computational Intelligence Synergies* *of Fuzzy Logic, Neural Networks and Evolutionary Computing* (1st ed.). New Delhi, India: John Wiley & Sons.

Syafaruddin, Karatepe, E., & Hiyama, T. (2011). Controlling of artificial neural network for fault diagnosis of photovoltaic array. *16th International Conference on Intelligent System Applications to Power Systems*, 1–6. https://doi.org/10.1109/ISAP.2011.6082219

Vieira, R. G., de Araújo, F. M. U., Dhimish, M., & Guerra, M. I. S. (2020). A Comprehensive Review on Bypass Diode Application on Photovoltaic Modules. *Energies*, *13*(10), 2472. https://doi.org/10.3390/en13102472

Vieira, R. G., Dhimish, M., de Araújo, F. M. U., & Guerra, M. I. S. (2020). PV Module Fault Detection Using Combined Artificial Neural Network and Sugeno Fuzzy Logic. *Electronics (Switzerland)*, *9*(12), 1–21. https://doi.org/10.3390/electronics9122150