An enhanced machine learning based approach for failures detection and diagnosis of PV systems

Elyes Garoudja[a], Aissa Chouder[b], Kamel Kara[a], Santiago Silvestre[c]

[a] SET Laboratory, Electronics Department, Blida 1 University, BP 270 Blida, Algeria
[b] Electrical Engineering Laboratory (LGE), University Mohamed Boudiaf of M'sila, BP 166, 28000, Algeria
[c] MNT Group, Electronic Engineering Department, Universitat Politécnica de Catalunya (UPC) BarcelonaTech, C/Jordi Girona 1-3, Campus Nord UPC, 08034 Barcelona,
Spain

Abstract:

In this paper, a novel procedure for fault detection and diagnosis in the direct current (DC) side of PV system, based on probabilistic neural network (PNN) classifier, is proposed. The suggested procedure consists of four main stages: (i) PV module parameters extraction, (ii) PV array simulation and experimental validation (iii) elaboration of a relevant database of both healthy and faulty operations, and (iv) network construction, training and testing. In the first stage, the unknown electrical parameters of the one diode model (ODM) are accurately identified using the best-so-far ABC algorithm. Then, based on these parameters the PV array is simulated and experimentally validated by using a PSIM™/Matlab™ co-simulation. Finally, efficient fault detection and diagnosis procedure based on PNN classifier is implemented. Four operating cases were tested in a grid connected PV system of 9.54 kWp: Healthy system, three modules short-circuited in one string, ten modules short-circuited in one string, and a string disconnected from the array. Moreover, the PNN method was compared, under real operating conditions, with the feed forward back-propagation Artificial Neural Network (ANN) classifiers method, for noiseless and noisy data to evaluate the suggested method's accuracy and test its aptitude to support noisy data. The obtained results have demonstrated the high efficiency of the proposed method to detect and diagnose DC side anomalies for both noiseless and noisy data cases.

Keywords: Photovoltaic, Fault Detection, Diagnosis, Probabilistic neural network, Artificial neural network, Best-so-far ABC, Maximum Power Point.

1. Introduction :
1.1. State of the art :

The amount of photovoltaic (PV) plants, around the world, has risen in a very remarkable way these last few decades. This is mainly due to the raised concerns for solving environmental pollution issues, harsh climate changes, and depletion of fossil energy sources [1]. Indeed, traditional energy sources, such as nuclear, oil and coal, have undesirable effects on ecosystem biodiversity, human health and climate change. In contrast, renewable energy sources, such as PV systems, are sustainable, safe and eco-friendly with zero CO_2 emissions [2–5]. An investigation study by Masson et al. [6] has revealed that more than 53 million tons of CO_2 emissions could be avoided in Europe in 2012 by producing 100 GW from PV systems.

PV systems, especially the DC side "PV array", are designed to operate under harsh outdoor conditions such as bad weather conditions, wind vibration effects on PV modules, ultraviolet radiation and corrosion [7]. Under such conditions, several malfunctions and failures may occur, resulting in PV modules lifetime reduction, lower total energy yield, loss of availability and safety hazards [8–12].

Due to the aforementioned considerations, fault detection and diagnosis of PV systems has seen a particular interest in the last few decades, where several research studies and investigation have addressed the issue [13]. These methods are mainly classified into two categories: conventional threshold approaches and machine learning based approaches [7]. In conventional threshold approaches, fault detection and diagnosis can be achieved by analyzing several electrical indicators, such as operating current and voltage, and the output generated power. For instance, Chouder et al. [14] suggested an efficient procedure for a supervision and fault detection in PV systems on the basis of power losses analysis. This method defines four new indicators to reach the fault detection and supervision goal: current ratio, voltage ratio, thermal capture losses and miscellaneous capture losses. A fault detection and remote monitoring strategy has been proposed by Drews et al. in [15]. This method uses irradiance values derived from satellite data instead of using real on-site measurements. The fault occurrence is declared when a significant deviation between the measured and simulated output power is detected. Garoudja et al. [2] proposed a novel fault detection and diagnosis strategy based on the exponentially weighed moving average (EWMA) monitoring chart. This method computes MPP residuals of current, voltage and power using actual measurements and the simulated fault-free data. The EWMA based method can detect three types of DC side faults such as: short circuit faults, open circuit faults and temporary shading faults.

Although threshold assessment approaches are straightforward and can yield reliable results, they still suffer from manual threshold calculations and require the photovoltaic array to operate at the MPP. The second category of fault detection and diagnosis in PV systems is machine learning based approaches. Actually, machine learning-based approaches require the collection and preprocessing of faulty and healthy data, from an operational PV system, to detect and diagnose eventual faults. For instance, a three-layered artificial neural network (ANN) based method has been proposed by Karatepe et al. in [16]. This method can only detect and localize short-circuit PV modules in a PV array. Zhao et al. [17–20] proposed several methods based on outlier detection rules, graph based semi supervised learning, and decision tree methods to detect and diagnose several faults in a PV system such as: open circuit faults, short circuit faults, line-line faults, partial shading occurrence and degradation fault. Outlier detection rules method requires the use of current and voltage measurements of each separate PV module, which makes it a very costly strategy. Additionally, decision tree method requires a lot of high-quality data to establish an accurate training stage.

Artificial neural network (ANN) has been also used to detect and classify eight categories of PV system failures in [21,22]. In these works, ANN has been combined with other threshold based methods to achieve fault detection and diagnosis in PV systems. The fault detection was carried out using a threshold approach while faults classification was reached by ANN classifier. In all these works, ANN networks are of multilayer perceptron (MLP) type, trained iteratively via gradient descendant algorithm to compute the back propagation (BP) difference between real outputs and targets. This type of network suffers from its slow training steps, and it could fall in local minima instead of global minima. These weaknesses could affect its reliability, efficiency and even its real-time implementation [7]. Moreover, this kind of network requires a lot of high-quality labeled data that describes very well the process [7].

Probabilistic neural network (PNN) [23,24] has been recognized as an efficient strategy to solve classification problems. It has several important advantages such its computation speed and simplicity during the training process [23]. This is mainly due to the no need to weights adaptations. Also, this network can classify any new incoming sample without repeating extensively the entire training process [23]. More particularly, PNN classifier has a strong robustness against noisy data generated by the measurement equipment and involved sensors. In the present work, a fault detection and diagnosis

of a grid connected PV system approach, based on PNN classifier, is proposed. In this approach, a trusted model of an actual operating PV systems is firstly established using the one diode model (ODM) and a parameters extraction procedure. After experimental validation of the PV generator model, healthy samples as well as faulty samples are generated and then used as inputs to the PNN classifier. This approach uses two PNN classifiers; the first one is used to detect faults occurrence, while the second classifier aims to diagnose the type of fault. The proposed method of fault detection and diagnosis consists of four main stages: (i) PV module parameters extraction based on the best-so-far ABC algorithm [25], (ii) Model validation, (iii) relevant dataset elaboration, and finally (iv) fault detection and diagnosis based on PNN classifiers. The main advantages of this approach are its high robustness against noisy data generated by measuring equipment and sensors. In fact, until recently, none of the proposed strategies tests the robustness of its method against noisy measurements.

This strategy has been tested based on real experimental data collected from 9.54 kWp Grid connected PV system, situated at the rooftop of the administrative building of CDER "Centre de Développement des Energies Renouvelables" in Algiers, Algeria. This method was tested under four operating cases: healthy system, three modules shortcircuited in one string, ten modules short-circuited in one string, and a string disconnected from the array. In addition, the proposed method based on PNN classifier was compared, under real operating conditions, with a feed forward back propagation artificial neural network (ANN) method. This comparison was established for noiseless and noisy data in order to test the proposed strategy's efficiency and robustness.

The present paper is arranged in the following manner. Section 2 outlines the description of the PV system and the faults detection and diagnosis strategy. Section 3 describes in detail the proposed detection and diagnosis strategy. In Section 4, simulations and experimental results are shown and analyzed in details. Finally, Section 5 is reserved to the drawn conclusions.

2. Description of the PV system and the faults detection and diagnosis strategy
2.1. PV description

PV plant under study is an actual Algerian grid-connected photovoltaic (GCPV) plant located in Algiers (Algeria) (Latitude: 36 °43′N,Longitude: 3 °15′E). This PV plant is a 9.54 kWp arranged in 3 subarrays of 3.18 kWp each one. Each sub-array contains 30 Isofoton 106W-12V PV modules mounted in 2 parallel strings of 15 PV modules in series and connected to a 2.5 kW (IG30 Fronius) single phase inverter.

Tilted and horizontal irradiances are measured using a Kipp & Zonen CM11 thermoelectric pyranometer, while PV module temperature is measured using K-type thermocouple. Meteorological variables as well as electrical variables measurements are performed by a data logger (Agilent 34970) as depicted in Fig. 1. Table 1 summarizes the main electrical parameters of Isofoton 106-12 PV module given under standard test conditions (STC) (Temperature level of 25 °C and irradiance level of 1000 W/m^2).
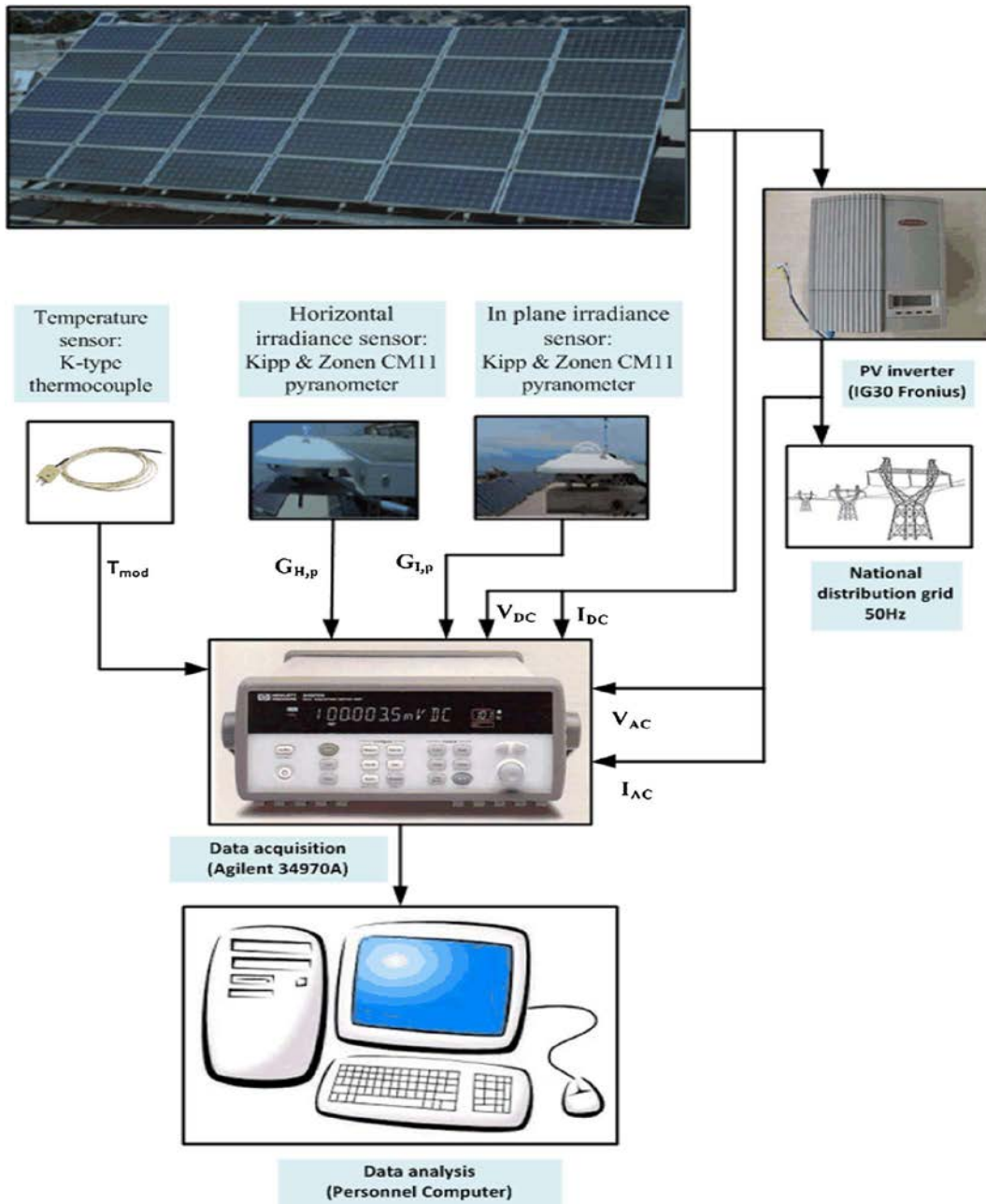
Fig. 1. PV plant under study and the monitoring system.

Table 1: Isofoton 106-12 PV module main electrical parameters under STC

Parameters	I_{mpp0} [A]	V_{mpp0} [V]	P_{mpp0} [W]	I_{sc} [A]	V_{oc} [V]
Values	6.4	17.4	106	6.54	21.6

2.2 Faults detection and diagnosis strategy

The main objective of this work is to design an efficient and reliable procedure, based on PNN classifier, to detect faults in a photovoltaic system and to diagnose their origin. However, using this

type of method, to deal with the classification problem, requires the availability of high-quality database that describes very well the process for each class. Practically speaking, obtaining such a database cannot often be guaranteed, especially in PV systems. In fact, operating a PV system under some types of failures can make the system completely insecure and cause catastrophic damages and safety hazards. Therefore, the best way to deal with this concern is to have a trusted simulation model that mimics the actual behavior of a PV system under healthy and faulty states. The flowchart, given in Fig. 2, summarizes the adopted steps followed to construct the fault detection and diagnosis strategy, namely: PV module parameters extraction, experimental validation, database elaboration and finally fault detection and diagnosis.

2.2.1. PV module parameters extraction

Usually, the solar cell/PV module is described by the well-known one diode model (ODM). This model describes the PV module behavior by the following electrical circuit (see Fig. 3):

In this model, the output current versus output voltage is given by the following equation [2,14]:

$$I_{PV} = I_{ph} - I_0\left(\exp\left(\frac{q(V_{PV} + R_s I_{PV})}{nk_B T}\right) - 1\right) - \underbrace{\frac{V_{PV} + R_s I_{PV}}{R_{sh}}}_{I_{sh}}$$

$$\underbrace{\qquad\qquad\qquad\qquad\qquad\qquad}_{I_d}$$

(1)

where: I_{PV} and V_{PV} are current and voltage outputs of the PV module respectively; Iph stands for the light-generated current. I_0 is the diode saturation current; Rs and Rsh are series and parallel resistances respectively; n is the diode ideality factors; kB is Boltzmann constant ($k_B = 1.3806503 \times 10^{-23}$ j/k) , T is the solar cell/PV module temperature and q is the electronic charge (q= $1.60217646 \times 10^{-19}$ c).

The ODM parameters extraction stage can be described as an optimization problem, where the cost criteria to be minimized is the Root Mean Square Error (RMSE) between the measured current and estimated current using Eq. (1), as given bellow:

$$RMSE = \sqrt{\frac{I}{S}\sum_{i=1}^{S}\left[g_i\left(I_{meas}, V_{meas}, \gamma\right)\right]^2} \qquad (2)$$

where:

$$g\left(I_{meas}, V_{meas}, \gamma\right) = I_{meas} - \left(I_{ph} - I_0\left(\exp\left(\frac{q\left(V_{meas} + R_s I_{meas}\right)}{nk_B T}\right) - 1\right) - \frac{V_{meas} + R_s I_{meas}}{R_{sh}}\right) \qquad (3)$$

I_{meas} and V_{meas} are the PV module output generated current and voltage respectively, and which are obtained from a set of experimental outdoor measurements using an IV curve tracer (PVPM40). The letter γ defines the vector of ODM parameters, and S is the size of experimental (I-V) curve data used during the ODM parameters extraction stage. ODM parameters extraction procedure is highlighted in Fig. 4. The used temperature during this step denotes PV module temperature. The main goal of this

stage is to find the optimal ODM parameters [I*ph,I*₀,n*,R*s,R*sh] that minimize the cost criteria and gives the lowest RMSE value.
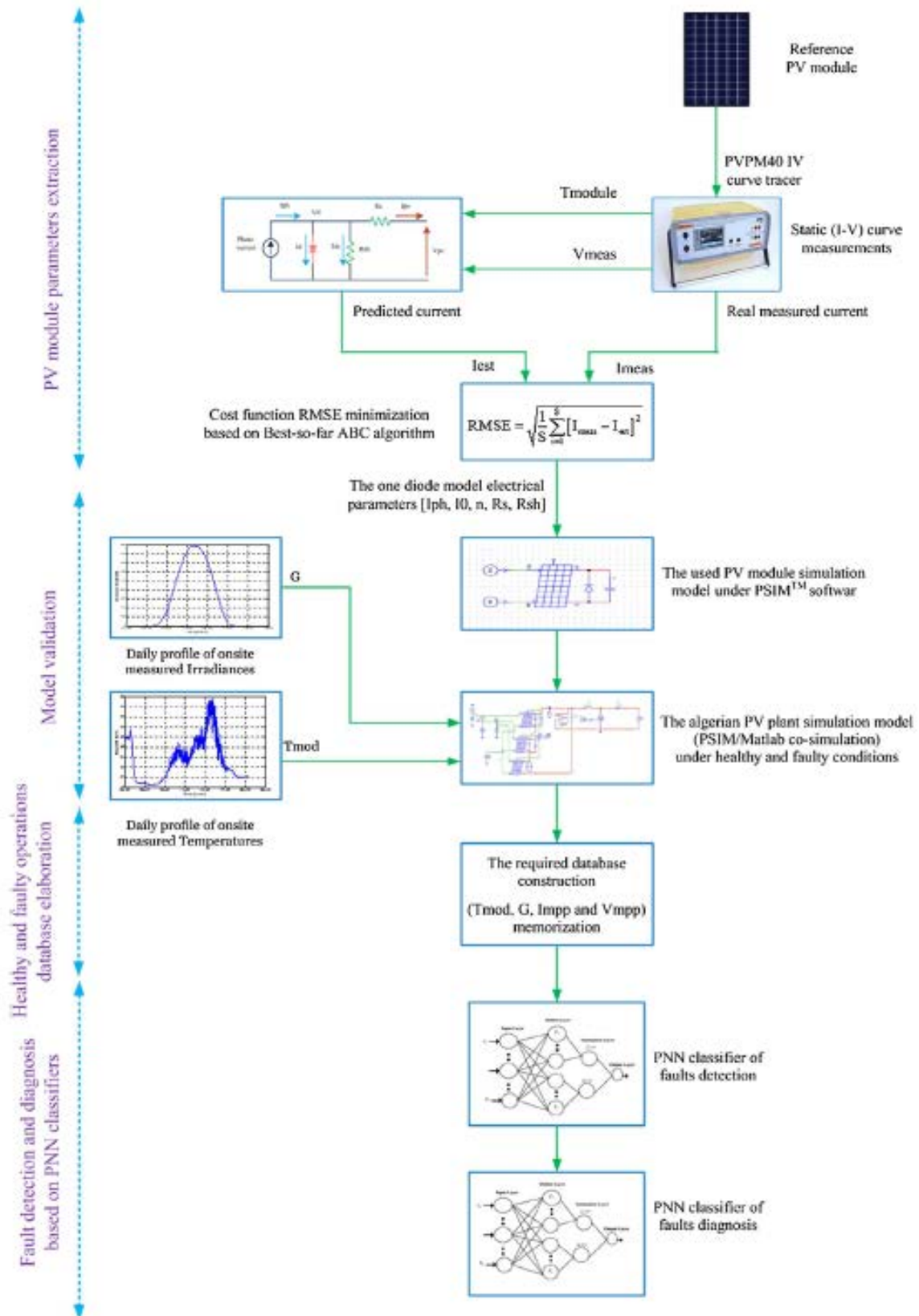

Fig. 2. Flowchart of the followed steps of the proposed faults detection and diagnosis strategy.
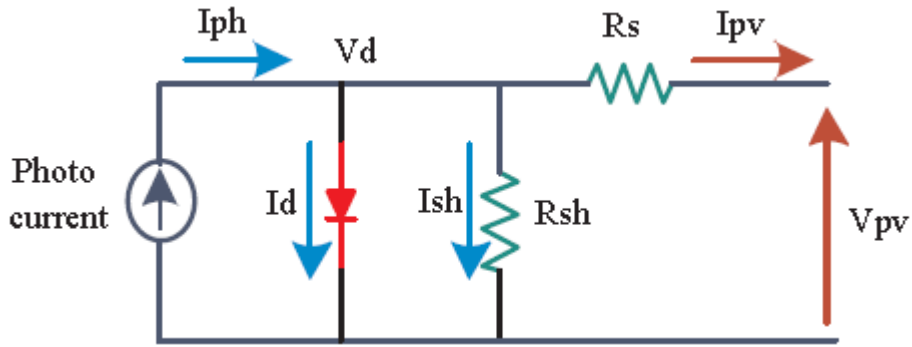
Fig. 3. Solar cell circuit model.

In this work, the use of the best-so-far ABC algorithm; an improved variant of the canonical ABC algorithm; to accurately extract the ODM five electrical parameters is proposed. The best-so-far ABC algorithm is efficient and accurate in solving optimization problems [26–28]. This is mainly due to its capacity to converge, regardless initial conditions, to accurate global solutions, high convergence speed and its simplicity in terms of implementation [25]. Moreover, in our previous works [2,29–31], the canonical ABC algorithm [32] was successfully applied to extract the ODM parameters.

2.2.2. Model validation

The identified ODM parameters are, subsequently, used to simulate the real PV system for normal conditions using PSIM™/Matlab™ Cosimulation. Indeed, the physical model of solar module included in PSIM™ software is used to simulate the physical behavior of the actual PV system. This can be reached by introducing the ODM identified parameters into this model, and then simulate the whole PV system for real daily profile measurements of temperatures and irradiances. On the other hand, data processing and involved calculations are performed in Matlab/Simulink™ environment. Finally, the simulated power (Psim) and the measured one (Pmeas) are then compared.

2.2.3. Database elaboration

This stage consists of elaborating a high-quality database that accurately describes the system behavior, in normal and faulty conditions. To this end, the validated PV system model is used to generate healthy and faulty samples, by intentionally introducing the desired faults, with real daily temperature and irradiance profiles. The recorded samples, corresponding to each operating condition, include four attributes: module Temperature "T", tilted Irradiance "G", Current at MPP "Impp" and Voltage at MPP "Vmpp".

2.2.4. Fault detection and diagnosis based on PNN classifier

Based on the elaborated data base, the last step consists of constructing two probabilistic neural network (PNN) classifiers: the first one is dedicated to fault detection and the second is responsible for diagnosing the origin of faults. The methodology of PNN construction, training and test are explained in details hereafter.
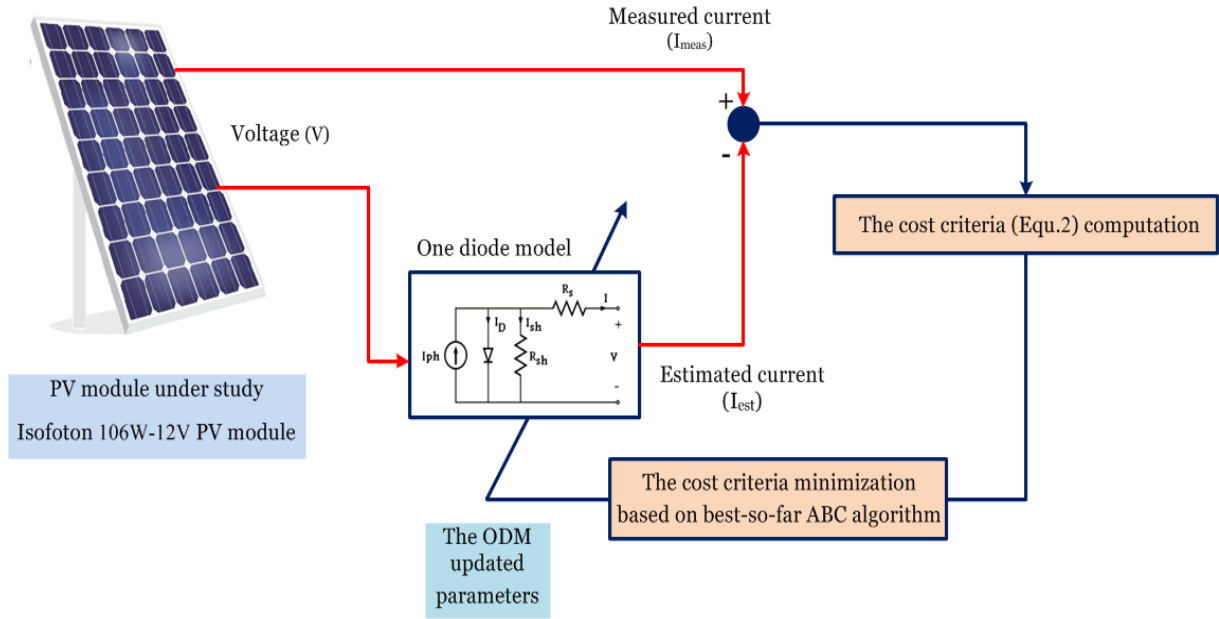
Figure 4: The ODM parameters identification procedure.

3   Detection and diagnosis strategy

3.1. Probabilistic neural network

Probabilistic neural network (PNN) [23,24], as a learning based method, has been considered as a powerful classification technique. Similarly to the other neural network strategies, the PNN uses a training set to extract pattern statistics, and a testing set to check the classification accuracy. Its architecture is similar to that of the back propagation one [33], except of using the exponential function, as an activation function, instead of using the sigmoid function [24]. As shown in Fig. 5, the probabilistic neural network consists of input units, pattern units, summation units and output units.

The aim of the input units is to directly distribute the input vector X to the pattern units, while the number of neurons in these units corresponds to the training set size. Besides, the pattern units, described by Fig. 6, consider the use of the same number of neurons as the input units. The pattern units are responsible for dot multiplication of the elements of the input vector and their respective weights $Z_i = X \cdot W_i$, given that the weight coefficient W_i of each pattern unit is set to the same value X_i of the training set. The result, Z_i, will be then introduced to a nonlinear activation function, given by the following expression:

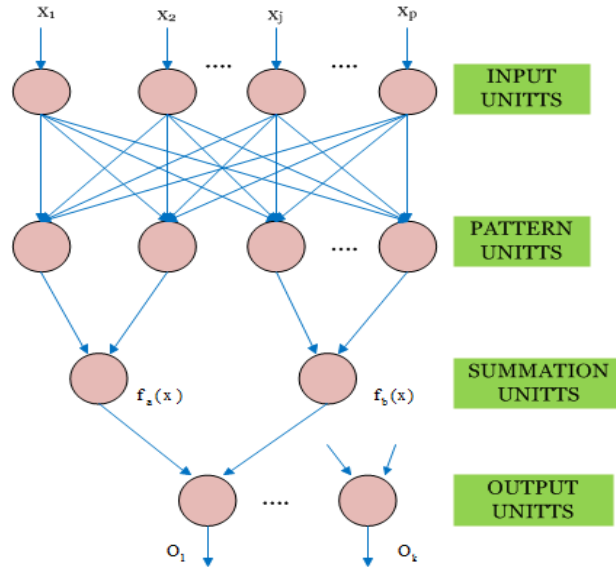$$Q(Z_i) = \exp\left(\frac{(Z_i - 1)}{\sigma^2}\right) \tag{4}$$

Figure 5. PNN structure for classification problem of two dimensions.

In the case of normalized input and weight vectors, the activation function (Equation 4) will result in:

$$Q(X) = \exp\left(\frac{-(W_i - X)^t (W_i - X)}{(2\sigma^2)}\right) \qquad (5)$$

where σ denotes the smoothing parameter.

The outputs of the pattern units will then be transmitted to the summation units, in which the number of neurons matches the number of classes. The outputs of each summation unit are given by:

$$f_A(X) = \frac{1}{(2\pi)^{\frac{P}{2}}\sigma^P}\frac{1}{m}\sum_{i=1}^{m}\exp\left[\frac{-(W_i - X_{Ai})^t (W_i - X_{Ai})}{2\sigma^2}\right] \qquad (6)$$

where: $f_A(X)$ is the Probability Density Function (PDF) of the input vector X; P is the number of patterns; m is the learning set size; X_{Ai} describes the corresponding i[th] training pattern of class A. The outputs units define the decision blocks. PNN comprises only one neuron in its output layer. This layer receives, from the summation layer, the probability density functions (PDF) values of each class and predicts the adequate class of the new sample. In other terms, this last layer outputs the adequate target that corresponds to the highest probability density function in summation units. The efficiency of PNN classifier depends on the used PDF accuracy, which is defined by the best choice of its smoothing parameter value [24].

Compared with conventional back-propagation network, PNN owns several important advantages [23,24] such as:

- Its training stage requires only one single pass (neither iterations nor weights computation).
- It can support both erroneous and noisy samples.
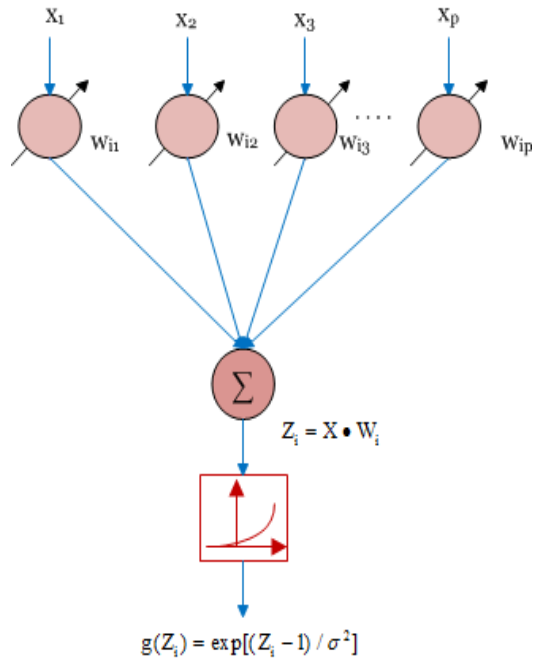- It has only one adjustable parameter (the smoothing parameter σ).

$$Z_i = X \bullet W_i$$

$$g(Z_i) = \exp[(Z_i - 1) / \sigma^2]$$

Figure 6. PNN Pattern unit.

3.2 *PNN based method for fault detection and diagnosis.*

As stated before, the objective of this work is to detect and diagnose faults in the DC side of a grid connected PV systems based on PNN classifier. Using PNN as fault detection and diagnosis approach requires four essential steps: elaboration of a relevant database that contains the necessary data to train and test the network, network construction, training the network using the learning set, and finally test and evaluation of the network efficiency based on the testing set.

3.2.1. Elaboration of the relevant database

To be able to obtain a pertinent database for both healthy and faulty systems, the physical model of the PV module implemented in PSIM™ software, is used to simulate, under healthy and faulty conditions, the real operation of the previously described Algerian grid connected PV system. The employed PV array comprises two parallel strings of fifteen PV modules in series (Isofoton 106W-12V). During this simulation, the physical model of the PV array requires the introduction of the five electrical parameters values, which are obtained using the identification algorithm described in Section 2.2. The proposed strategy for simulating the photovoltaic system using the PSIM™ software is characterized by its physical nature, which makes it possible to easily simulate the behavior of several faults that usually happen in the DC side of PV systems, such as: shading faults, short circuit faults, open circuit faults and line-line faults, …etc. Moreover, the used simulation model requires much reduced time to simulate the behavior of PV system whatever its physical configuration, and even for a very high number of meteorological data.

In this paper, the simulations of four operating cases have been considered: healthy system; three modules short-circuited in a string; ten modules short-circuited in a string, and finally a completely disconnected string in an array, as depicted in Fig. 7. To finish with, the final database is constructed considering four attributes [T, G, Impp,and Vmpp] for each operating case. Fig. 8 shows the flowchart of PNN dataset elaboration.

3.2.2. The networks construction

Since the proposed work aims both fault detection and diagnosis, two PNN classifiers are then constructed. Indeed, the occurrence of any fault in the PV system is detected by the first network, while the diagnosis of its type is carried out by the second one. The detection network, shown in Fig. 9, has four inputs (T, G, Impp and Vmpp) and two outputs (healthy state, faulty state), while the diagnosis network, shown in Fig. 10, has the same inputs as the first one, and three outputs (fault #1, fault #2, fault #3) corresponding to the three faulty operating cases. PNN classifiers of fault detection and diagnosis (Figs. 9 and 10) comprise only one neuron in their output layers. This neuron outputs the class that corresponds to the highest probability density function in summation units.

In order to avoid any conflict, the detection and diagnostic networks are connected in series so that the diagnostic network will not be activated as long as the faulty alarm does not occur from the detection network. This concept is explicitly explained by the flowchart of Fig. 11, given below.

3.2.3. Learning stage

Now, the detection and diagnostic networks are formed using the learning data set previously obtained. It should be noted here that, on one hand, the input data (T, G, Impp and Vmpp), are of continuous type which are directly processed by the two PNNs. On the other hand, the outputs (healthy state, faulty state, fault #1, fault #2 and fault #3) are nominal variables that are not supported by this type of networks. For this reason, particular codes, arbitrary chosen, have been assigned to the PNN outputs as mentioned below:

• Healthy system: code "2", and which will be used for faults detection network to describe the healthy operating case.

• Faulty system: code "8", and which will be used for faults detection network to describe the occurrence of possible DC side faults.

• Fault #1: code "4", and which will be used for the diagnosis network to describe the occurrence of short circuit of three PV modules in one string.

• Fault #2: code "6", and which will be used for the diagnosis network to describe the occurrence of short circuit of ten PV modules in one string.

• Fault #3: code "9", and which will be used for the diagnosis network to describe the occurrence of a complete string disconnection from the PV array.

Finally, to train detection and diagnosis networks, learning set values and the codes of their corresponding classes are used as input and target vectors respectively.
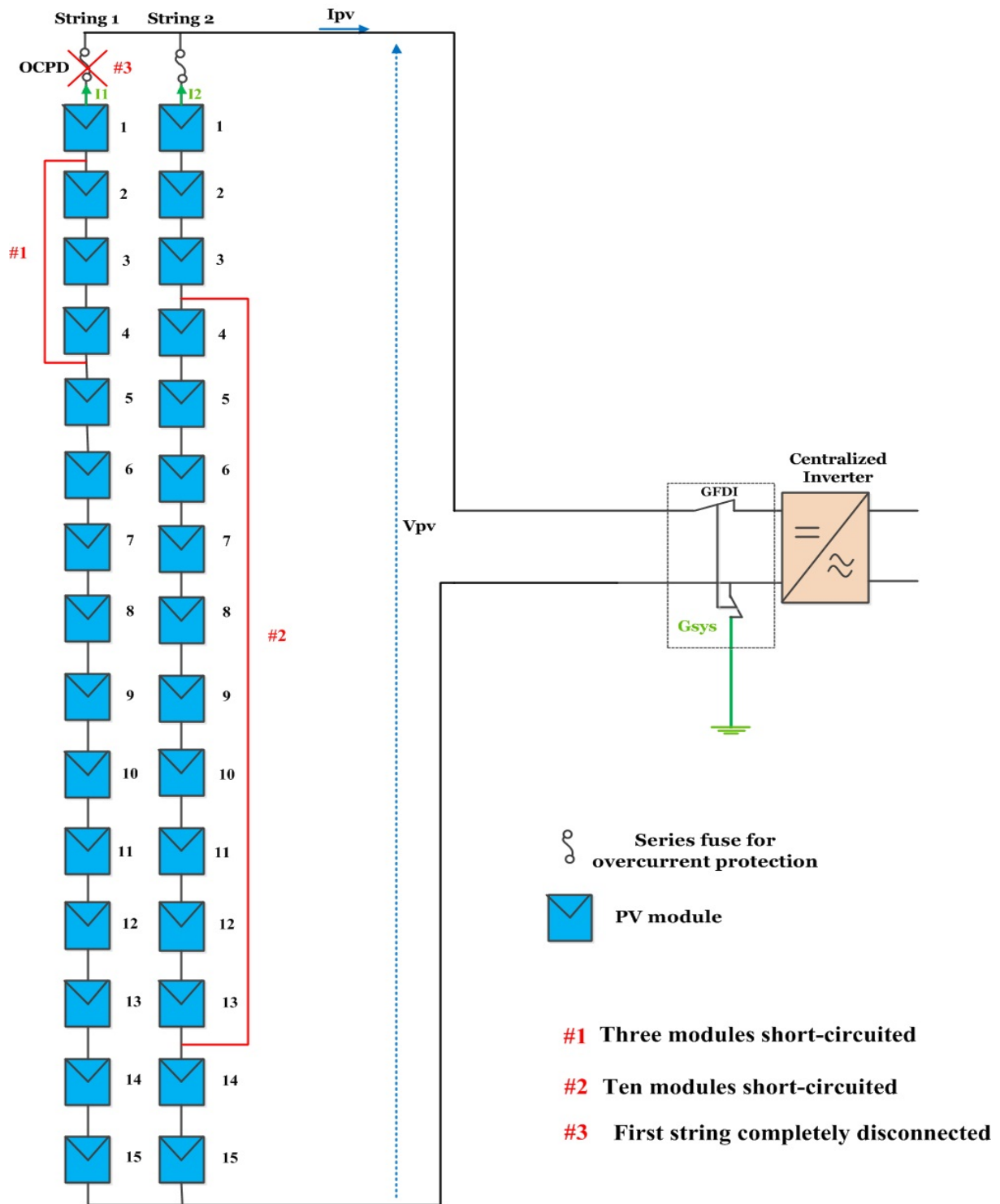
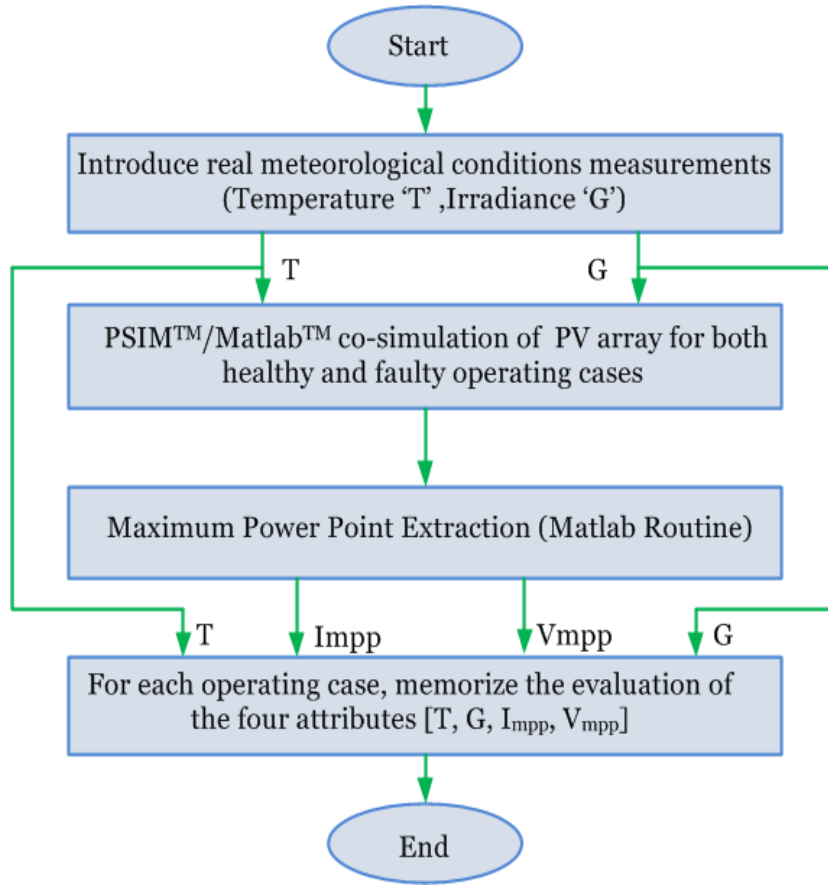Figure 7. The tested open circuit and short circuit failures of the Algerian PV system.

Figure 8. The database construction flowchart

3.2.4. Testing stage

After having learned both detection and diagnosis networks, the next stage is to test their effectiveness while using the testing set. Similarly, to the training stage, the input testing set to the two PNN classifiers are the four attributes (T, G, Impp and Vmpp), while their outputs represent their corresponding estimated classes.

In order to assess the networks efficiency, a performance test of the classifiers has been carried out using the following four standard metrics [34]:

$$\text{System Accuracy} = 100 \times \frac{\text{TP}_{all-classes}}{d} \qquad (7)$$

$$\text{Sensitivity(class(x))} = \frac{\text{TP}_{class(x)}}{\text{TP}_{class(x)} + \text{FN}_{class(x)}} \qquad (8)$$

$$\text{Specificity(class(x))} = \frac{\text{TN}_{class(x)}}{\text{TN}_{class(x)} + \text{FP}_{class(x)}} \qquad (9)$$

$$\text{Positive Predictivity(class(x))} = 100 \times \frac{\text{TP}_{class(x)}}{\text{TP}_{class(x)} + \text{FP}_{class(x)}} \qquad (10)$$

where:

TP is the True Positive classifications number, which denotes the number of samples which are supposed to be classified in class "x" and they are really classified into this class. FN is the False Negative classifications number, which denotes the number of samples from class "x" and are classified into another class than class "x". TN is the True Negative classifications number, which denotes the number of samples which are not supposed to be classified in class "x" and are really classified, according to the classifier, into another class than class "x". FP is the False Positive classifications number, which denotes the number of samples which are not supposed to be classified into class "x", while they really are classified, according to the classifiers, into this class, and 'd' defines the size of the testing set.

In addition to these performance evaluation standard metrics, the confusion matrix is another metric usually used to evaluate classifier performances. In fact, confusion matrix (also called contingency table) is a clean and unambiguous way to depict the classification results. As an example of a binary classification problem (Problem with two possible classes), the confusion matrix is composed of two rows and two columns, as can be seen in table 2.
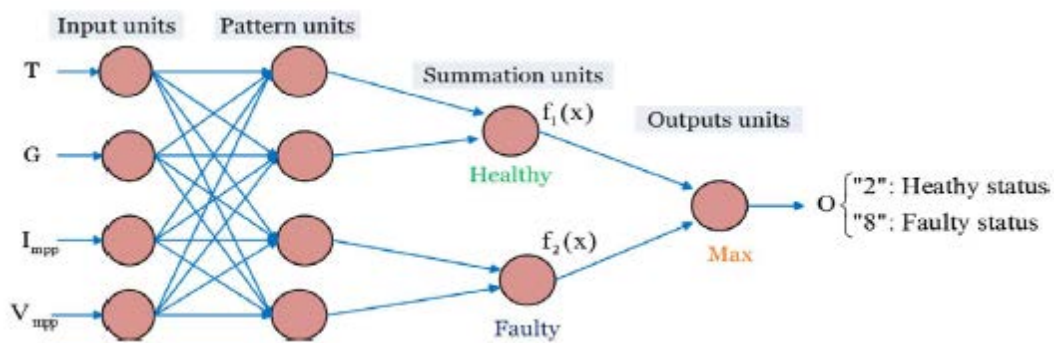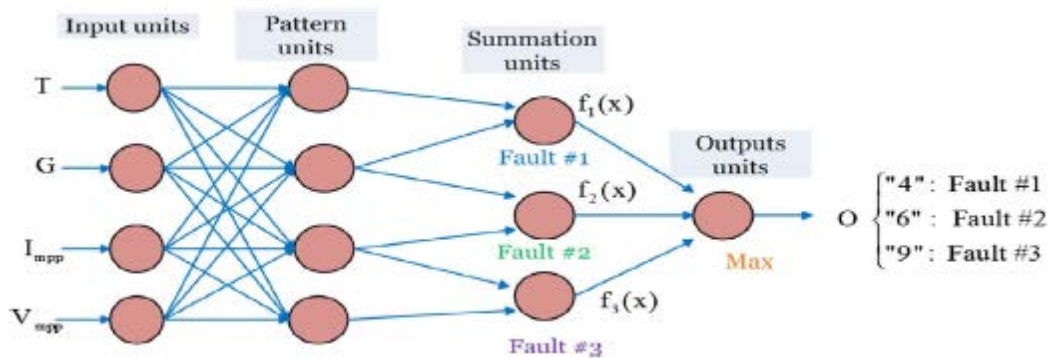

Fig. 9. PNN detection network.

Fig. 10. PNN diagnosis network.

Table 2

The confusion matrix of a binary classification problem.

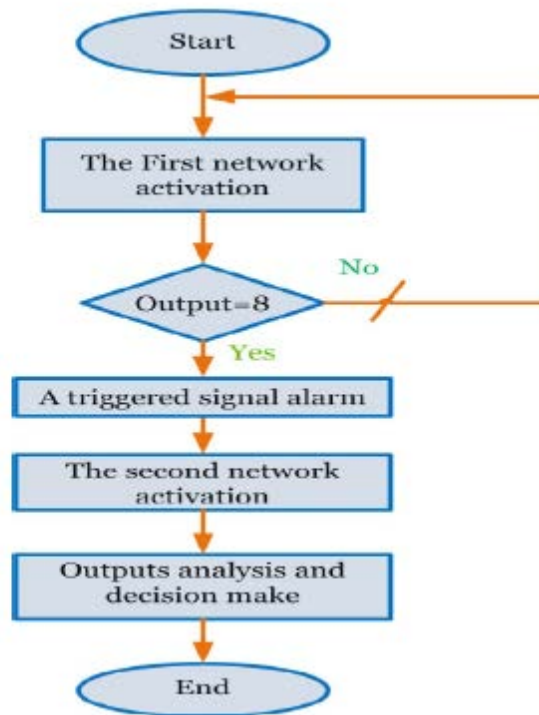|	The first class (Predicted)	The second class (Predicted)
The first class (Actual)	True Positive (TP)	False Negative (FN)
The second class (Actual)	False Positive (FP)	True Negative (TN)

Fig. 11. Flowchart of the two classifiers' connection.

4 Simulation and experimenta results

In order to verify the performance and effectiveness of the proposed approach, described previously in Section 2, a simulation study and an experimental evaluation are carried out, and the main results are highlighted. This section is subdivided into three sub-sections: PV module parameters extraction results, model validation results, and finally, PNN based fault detection and diagnosis results.

4.1. PV module parameters extraction results

In the first stage, the efficiency of PV module parameters extraction method based on the best-so-far ABC algorithm was validated experimentally.

To this end, experimental (I-V) curve measurements, using an I-V curve tracer (PVPM40), were recorded and compared with the prediction model (Eq. (1)). Module temperature and the irradiance level used in this particular test were 27.2 °C and 755 W/m^2 respectively.

The best-so-far ABC algorithm adjustable parameters are summarized in Table 3, while optimal values of the electrical parameters [Iph, I_0, n, Rs, Rsh], determined by the best-so-far ABC algorithm, are summarized in Table 4.

Table 3
The adjustable parameters of best-so-far ABC algorithm.

The adjustable parameters	Food source number	Number of employed bees	Number of onlooker bees	The solution dimension	Maximum cycles number
The selected value	150	75	75	5	2000

Table 4
Isofoton 106W-12V PV module's extracted parameters.

The electrical parameter	I_{ph} [A]	I_0 [A]	n	R_s [Ω]	R_{sh} [Ω]	RMSE
The identified value	5.54	1.11e 05	1.55	0.1474	202.5	0.014

The simulation result of the prediction model (Eq. (1)), obtained with the optimal parameters, and the experimental curves are shown in Fig. 12. It is clearly shown the good agreement between the experimental curve and the predicted one when simulated with the extracted parameters. Also, the smallest value of the RMSE, given in table 4, reveals the high accuracy of the best-so-far ABC algorithm to extract the best fitting parameters.
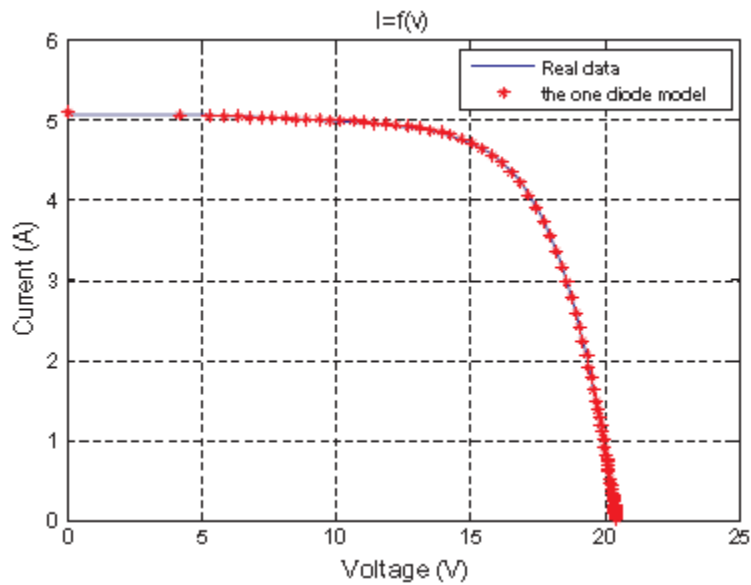

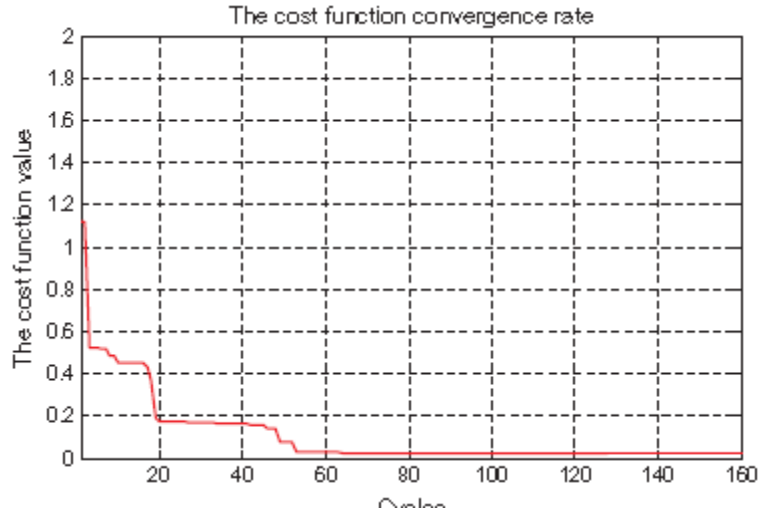Fig. 12. Real measured I-V curve against the estimated one.

Fig. 13. Best-so-far ABC algorithm convergence rate.

In Fig. 13, it is shown the best-so-far ABC algorithm convergence rate for each cycle until reaching the convergence condition (smallest RMSE value). While in Fig. 14, it is depicted the absolute error between measured current (Imeas) and estimated current (Iest), computed by the following equation.

$$Gap_{current} = |I_{meas} - I_{est}|$$

(11)

According to Fig. 14, the highest value of the absolute error does not exceed 0.06 which confirm the high accuracy of the best-so-far ABC optimization based algorithm to extract the best fitting parameters of ODM.
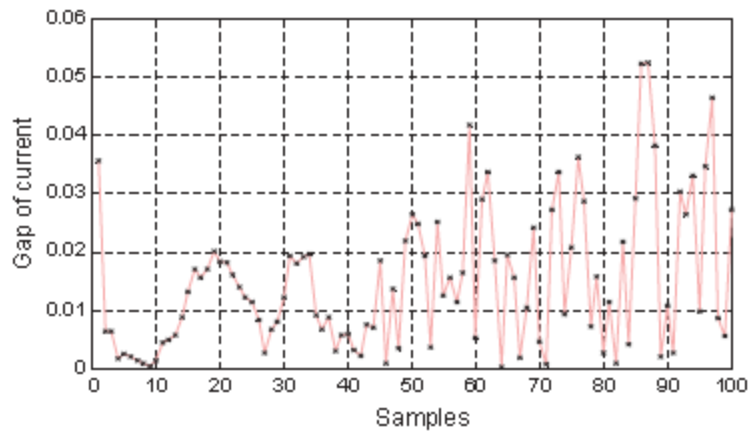

Fig. 14. Absolute error evolution between measured and estimated current.

4.2 Model validation results

The extracted parameters, calculated in the previous section, are now introduced into the whole PV plant that models the actual grid connected PV system located in CDER (Algiers), in normal operating conditions. The inputs to this model are the actual measured daily profiles of temperature and irradiance. In this work, a co-simulation methodology was adopted combining PSIM™ and Matlab™

software, where the physical system (PV generator and MPPT) is implemented in PSIM while data processing and the involved calculations are implemented in Matlab/Simulink™. The simulated output power (Pmppsim)is then compared to the output measured power (Pmpp-mea), under the same temperature and irradiance inputs, as shown in Fig. 15.
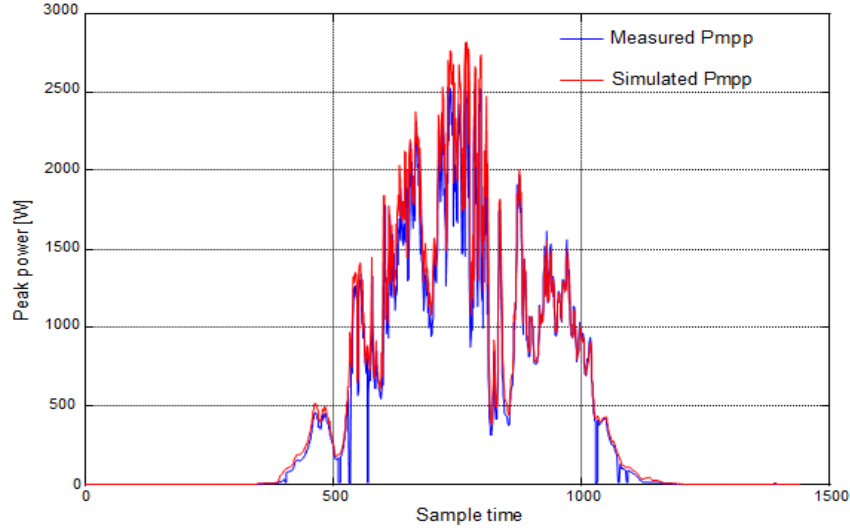

Fig. 15. Real measured against simulated peak power.

As this stage is crucial for the fault detection and diagnosis stage, detailed in the next section, the results given in Fig. 15 must be quantified in terms of discrepancy between measured and simulated powers. To this end, the Mean Absolute Error (MAE), given by Eq. (12), was considered to evaluate the daily mean gap between Pmpp-sim and Pmpp-meas. The calculated MAE for this particular day, characterized by fast-moving clouds, is found to be about 8 W.

$$\mathrm{MAE}_{Pmpp} = \frac{1}{L} \sum_{d=1}^{L} |P_{mppmeas} - P_{mppsim}|$$

(12)

where: L denotes the length of MPP power vector.

Table 5
ANN and PNN networks adjustable parameters.

	Feed forward back-propagation network (ANN)	Probabilistic neural network (PNN)
Detection network	The weights adjustment's method: back propagation algorithm. The number of hidden layer 25. The activation function: tangent sigmoid. The training phase's maximum number of epochs 100. The training phase's performance goal 10^{-5}.	The smoothing parameter σ 1.1. The activation function: exponential
Diagnosis network	The weights adjustment's method: back propagation algorithm. The number of hidden layer 40. The activation function: tangent sigmoid. The training phase's maximum number of epochs 100. The training phase's performance goal 10^{-5}.	The smoothing parameter σ 1.1. The activation function: exponential

Table 6
Classification accuracy and sensitivity results for noiseless data.

Detection networks tested with noiseless data

The classifier type	Classification accuracy (%)	Sensitivity (%)	
		Healthy operating case	Faulty operating case
ANN	90.35	100	78.14
PNN	100	100	100

Diagnosis networks tested with noiseless data

The classifier type	Classification accuracy (%)	Sensitivity (%)		
		Fault 1	Fault 2	Fault
ANN	73.18	86.96	100	32.61
PNN	100	100	100	100

Table 7
Confusion matrixes of ANN networks for noiseless data.

Detection network type ANN tested on the noiseless data

The real classes	The predicted classes	
	Healthy case	Faulty case
Healthy case	184	0
Faulty case	71	481

Diagnosis network type ANN tested on the noiseless test data

The real classes	The predicted classes		
	Fault 1	Fault 2	Fault 3
Fault 1	160	24	0
Fault 2	0	184	0
Fault 3	119	5	60

Table 8
Confusion matrixes of PNN networks for noiseless data.

Detection network type PNN tested on noiseless data

The real classes	The predicted classes	
	Healthy case	Faulty case
Healthy case	184	0
Faulty case	0	552

Diagnosis network type PNN tested on the noiseless test data

The real classes	The predicted classes		
	Fault 1	Fault 2	Fault 3
Fault 1	184	0	0
Fault 2	0	184	0
Fault 3	0	0	184

4.3. Fault detection and diagnosis results

In this subsection, the effectiveness of the proposed fault detection and diagnosis approach is presented in detail. At first, the PNN construction, training and validation are conducted with noiseless data, obtained from the simulation model, in normal and faulty operation. In the second step, the constructed PNNs are tested against noisy data, obtained by adding a predefined noise to the simulated data. In addition, a comparison study with classical ANN classifier for both noiseless and noisy data is also carried out in this subsection.

4.3.1. Noiseless data case

Once the simulation model has been validated, the previously described DC side failures have been simulated under measured daily profiles of temperature and irradiance. This simulation aims the elaboration of a high-quality database for both healthy and faulty systems as described in Section 3. The formed database contains 2960 data for each operating case, which means a total of 11840 data for the four simulated scenarios. For the training phase, we have chosen 75% (8880 data) while the test set contains the remaining 25% (2960 data). The details of the total database construction are given bellow:

• 740 data are associated with each of the four attributes (T, G, Impp and Vmpp) for each operating case. Thus, a total of 2960 =740 × 4 data set is associated with each operating case. More precisely, the vector of each attribute is subdivided into two sub-vectors: a first one comprises 556 samples (75% of 740) and it is stated as the classifier training set, while the second sub-vector includes the remaining 184 samples (25% of 740), and it is used as the classifier testing set.

• The test dataset comprises 184× 4= 736 data for each attribute, organized in the following manner:

The test set of healthy system includes 184 samples.

The test set of fault #1 includes 184 samples.

The test set of fault #2 includes 184 samples.

The test set of fault #3 includes 184 samples.

• On the other hand, this paper proposes the use of two classifiers (two different networks). The first PNN is responsible of failures detection, while the second classifier aims to diagnose accurately the type of the occurred fault. Therefore, the performance of the detection network is assessed using the test set in the following manner:
Healthy system test set includes 184 samples for each attribute.

Faulty system test set includes $184 \times 3 = 552$ samples (Denoted by the three faulty cases: Fault #1, Fault #2 and Fault #3) for each attribute.

• While the performance of diagnosis network is assessed using the test set in the following manner:

Fault #1 test set includes 184 samples for each attribute.

Fault #2 test set includes 184 samples for each attribute.

Fault #3 test set includes 184 samples for each attribute.

It is worth mentioning here that the measured data from the real operating system are considered as noiseless data as the quantification of error measurements and sensor drifts is a very difficult task. For comparison purpose, the performances of the proposed fault detection and diagnosis in PV systems are compared with those abstained with the classical ANN classifier. Toward this end, two ANNs have been constructed, trained and tested with the same dataset as the PNN. The number of the hidden layers in the ANN networks and the smoothing parameter, σ, for the PNN networks are obtained by trial and error tests. The main parameters used in this work, for each network, are summarized in Table 5. Classifiers efficiency (total efficiency) and classification sensitivity results for ANN and PNN classifiers are summarized in Table 6. While the confusion matrixes of the two networks (ANN and PNN for detection and diagnosis) are summarized in Table 7 and Table 8 respectively.

For more clarity of the obtained classification results, the confusion matrixes have been displayed graphically for the two networks (PNN and ANN). Figs. 16 and 17 show the simulation results of the ANN classifiers in detection and diagnosis stages respectively, while Figs. 18 and 19 show the results of the PNN classifiers.

The simulation results in terms of Specificity and Positive Predictivity for noiseless data are summarized in Tables 9 (ANN networks) and 10 (PNN networks) respectively.

According to the results given above, the high classification abilities of the proposed method, based on PNN classifiers, are clearly demonstrated, whether at the detection stage or at the diagnosis stage. In fact, all the evaluated metrics (classification accuracy, sensitivity, specificity and positive predictivity) give a 100% success rate for the classification of the input samples in their corresponding classes (see Table 6 and 10).

Table 9. Classification Specificity and Positive Predictivity for ANN networks under noiseless data

Detection network type ANN tested on the noiseless test data		
classes	Specificity (%)	Positive Predictivity (%)
Healthy case	87.14	72.16
Faulty case	100	100
Diagnosis network type ANN tested on the noiseless test data		
classes	Specificity (%)	Positive Predictivity (%)
Fault1	67.66	57.35
Fault2	92.12	86.38
Fault3	100	100

Table 10. Classification Specificity and Positive Predictivity for PNN networks under noiseless data

Detection network type PNN tested on the noiseless test data		
classes	Specificity (%)	Positive Predictivity (%)
Healthy case	100	100
Faulty case	100	100
Diagnosis network type PNN tested on the noiseless test data		
classes	Specificity (%)	Positive Predictivity (%)
Fault1	100	100
Fault2	100	100
Fault3	100	100

Furthermore, these high performances, to correctly classify the incoming samples in their right classes, are highlighted by the confusion matrices given in Table 8 and represented graphically in Figs. 18 and 19. However, the evaluation metrics for ANN classifiers (Table 6) have given lower percentage rate in both detection and diagnosis stages. For instance, a very low percentage rate is obtained for sensitivity metric (32.61%) in case of fault#3. This luck to classify some samples in their right classes, observed in ANN classifiers, is clearly observed in the confusion matrixes plots where data belonging to certain classes are wrongly classified in another classes (Figs. 16 and 17). Thus, with the case of noiseless data, the PNN classifiers have given better results than ANN classifiers regarding fault detection and diagnosis.
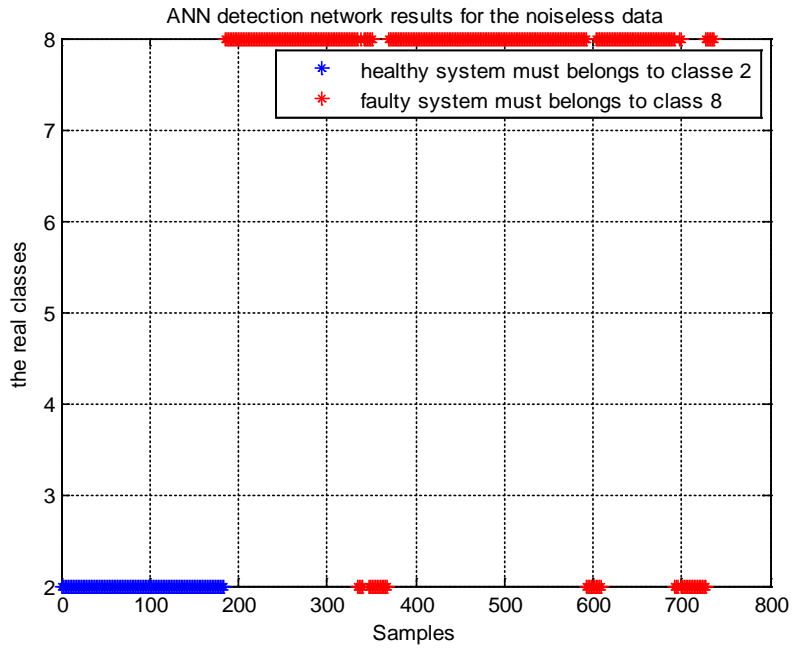
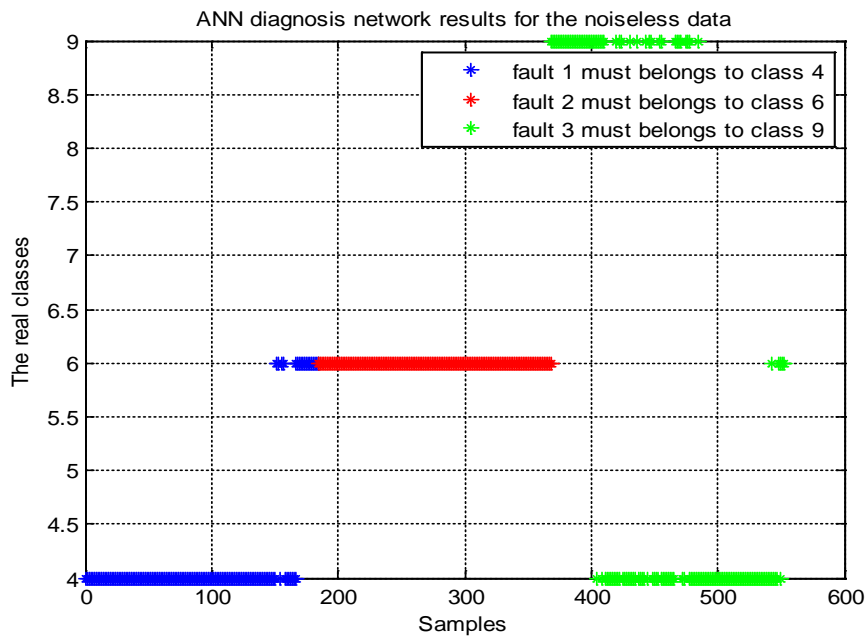Figure 16. Fault detection results based ANN classifier for the case of noiseless data.

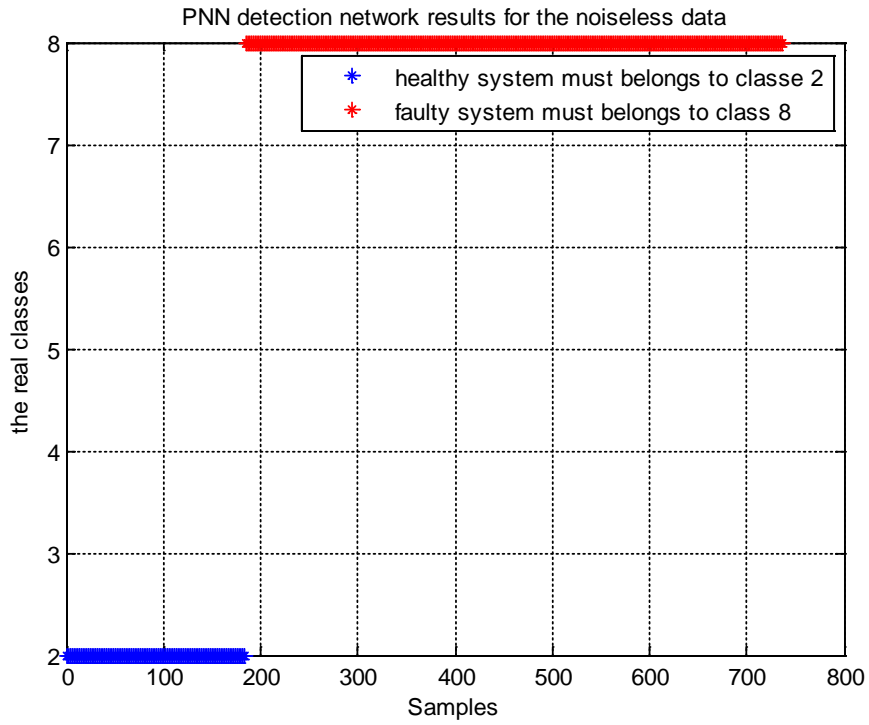Figure 17. Fault diagnosis results based ANN classifier for the case of noiseless data.

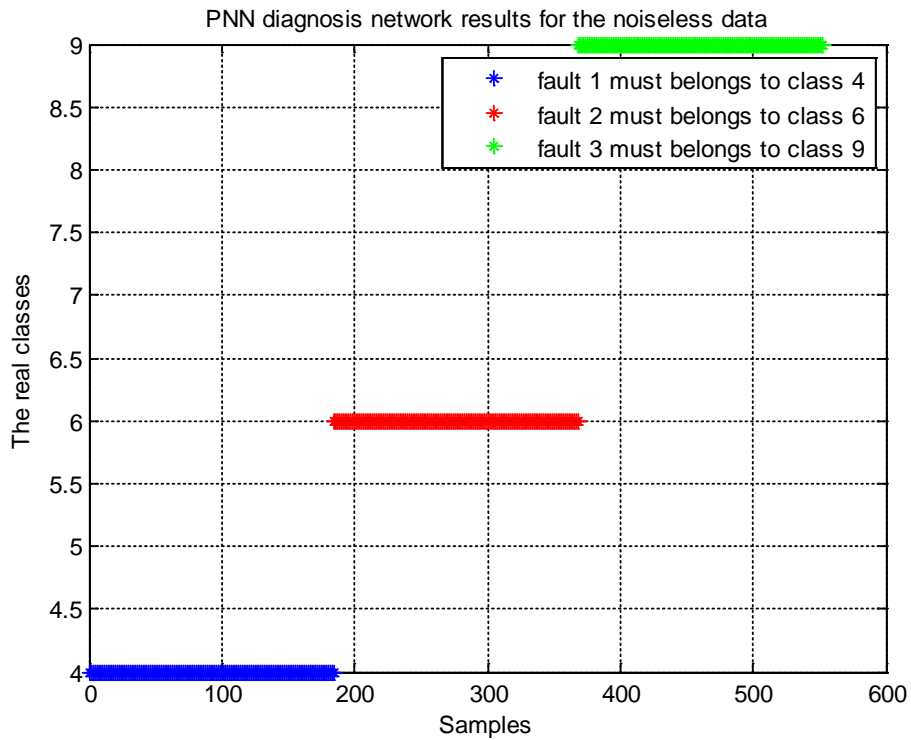Figure 18. Fault detection results based PNN classifier for the case of noiseless data.

Figure 19. Fault detection results based PNN classifier for the case of noiseless data.

4.3.2. Noisy data case

In practical cases, the measured quantities from the PV system are subjected to various types of disturbances and noises caused mainly by the tolerances of the measuring devices and sensor drifts. However, since these disturbances are difficult to quantify from a practical point of view and in order to test the performances of the proposed PNN with noisy data, we have considered the measured data as noiseless data where a predefined noise function is added to each attribute to form noisy data. This predefined noise function is a simple Matlab™ routine given by the following expression:

$$\text{Disturbance} = \alpha + \beta \times \text{randn}(1, N) \qquad (13)$$

where

α is the mean of the disturbance signal; β is the standard deviation of the disturbance signal and which defines its magnitude; "randn" is an implemented Matlab function used to generate a normal distributed number, and N the size of the disturbance signal. Subsequently, the two networks (ANN and PNN), previously constructed in Section 3.2.2 are tested on noisy data as follow:

 The disturbance signal is added to the corresponding attribute (T, G, Impp and Vmpp) using Eq. (13).

 The added noisy signals are characterized by their means equal to zero and the standard deviations equal to the disturbance magnitude.

 The chosen magnitudes of each disturbed attribute are:

$\beta_T = 4\ °C$, $\beta_G = 5\ W/m^2$, $\beta_{Impp} = 2\ A$, and $\beta_{Vmpp} = 5\ V$.

In order to test the networks efficiency with noisy data, the used learning set is composed of noiseless data, while the test set includes only the noisy attributes. By following the same assessment steps, described in Section 4.3.1, the simulation results in terms of Classification Accuracy (Total Efficiency), Classification Sensitivity, Confusion Matrixes, Specificity and Positive Predictivity for the case of noisy data are summarized in Tables 11–15 and graphically represented in Figs. 20–23.
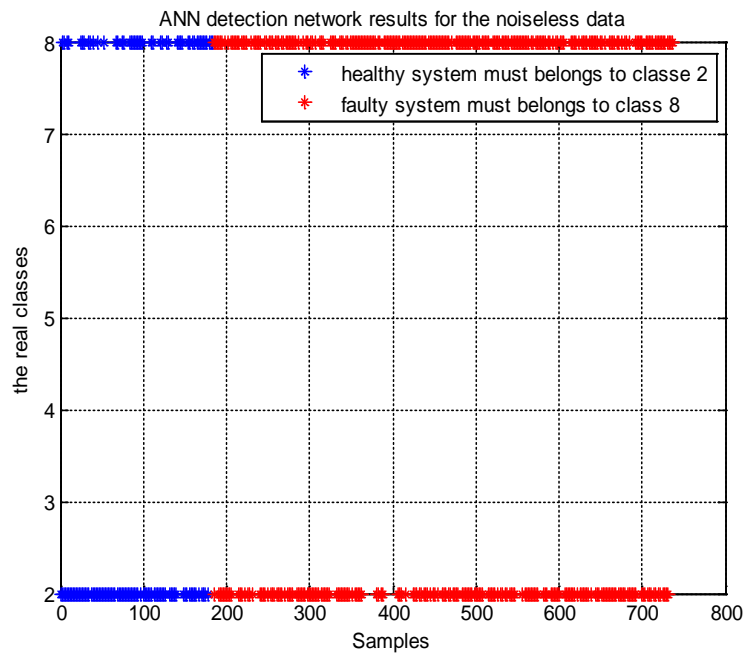
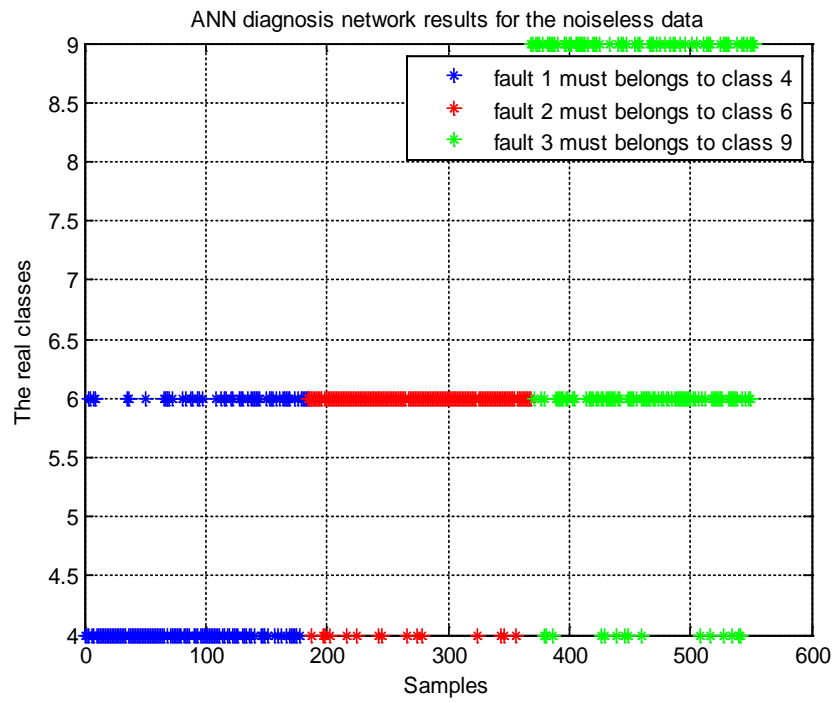Figure 20.Fault detection results based ANN classifier for the case of noisy data.

Figure 21. Fault diagnosis results based ANN classifier for the case of noisy data
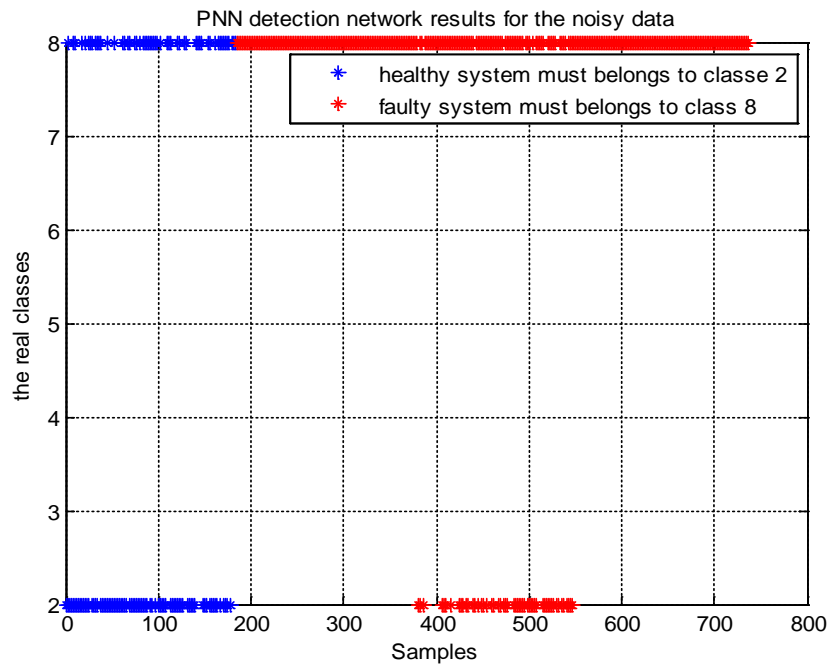

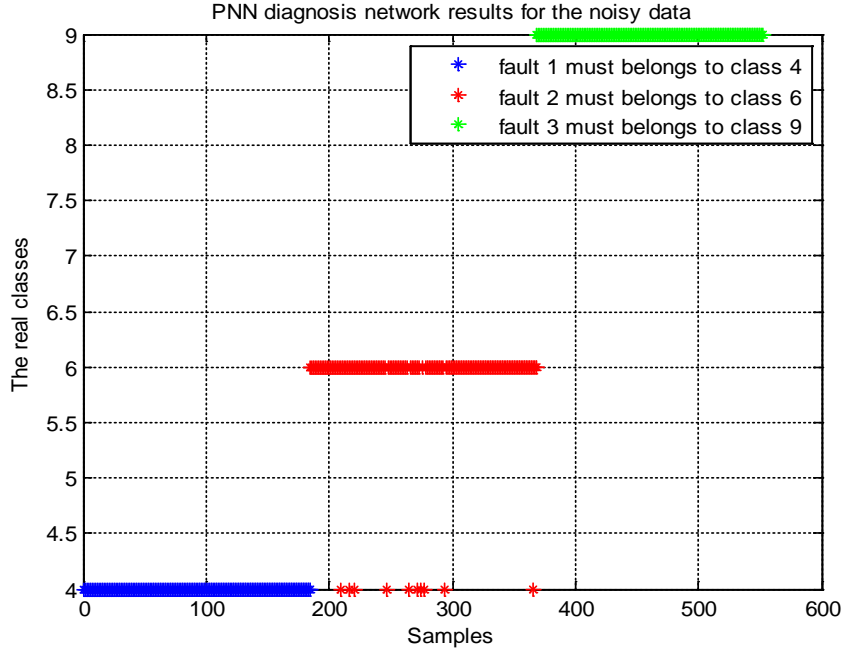Figure 22. Fault detection results based PNN classifier for the case of noisy data.

Figure 23. Fault diagnosis results based PNN classifier for the case of noisy data.

By closely analyzing the above results, it can be noted that even the exaggerated magnitudes of the noisy attributes (β_T = 4 °C, βG =5 W/m^2, β_{Impp} = 2 A, and β_{Vmpp} = 5 V), the two PNN classifiers still have a motivated results in both detection and diagnosis stages. In addition, according to Tables 11 and 15, the main performance metrics have not undergone a serious decline even under this test condition. For instance, the recorded classification efficiency, for the proposed classifiers, is 82.34% for detection network, and 98.19% for diagnosis network. In contrast, the test of the ANN classifiers under noisy data has led to a significant decrease of almost all the performance metric indicators. For instance, the recorded classification efficiency is 58.02% for detection network. Moreover, a very low percentage rate is obtained for positive predictivity metric (33.15%) in case of healthy state (Table 14). This significant decrease of the performances of ANN classifiers is also clearly observed in the confusion matrixes plots where several amounts of data belonging to certain class are wrongly classified in another classes (Figs. 20 and 21).

4.3.3. PNN classifiers' efficiency under real operating conditions

4.3.3.1. PNN classifier' efficiency under noiseless data.

As mentioned in the previous sections, the on-site measurements have been considered

as noiseless data in this work. Thus, in this subsection a real measured daily profile of temperature and irradiance is used to assess the detection and diagnosis capabilities of the proposed PNN classifier when the prescribed faults occur in a particular time of day. For clarity purposes, the measurements of a clear day are taken as a test pattern, where faults are intentionally introduced in the following manner:

From 2 h:00 to 8 h:00, the PV system is in normal working conditions (no faults), then a short-circuit of three PV module (Fault#1) in the same string has been introduced at 8 h:01. The system remains in that faulty condition until 9 h:30. At this time, the fault is cleared and the system comes back to its normal state. At 11 h:00, the system was forced to work under a condition of ten (10) short-circuited modules (Fault#2). The system keep working in this state until 12 h:30 where the fault is cleared and the PV system come back to its normal condition. After then, from 14 h:00 to 15 h:30 the first string of the PV array is completely disconnected (Fault#3). Thereafter, this fault is removed and the system continues working in normal conditions. These details are graphically shown in Figs. 24 and 25, for the monitored irradiance and module temperature, and Figs. 26, 27 and 28, for Impp, Vmpp and Pmpp.

The main results, obtained when using the previous patterns, are shown in Fig. 29 (detection) and in Fig. 30 (diagnosis). It is clearly shown that, before the occurrence of any fault, the detection PNN network (Fig. 29) generates 'Class 2', indicating healthy operating state While the diagnosis PNN generates, in this particular case, "class 0" which denotes that is not yet activated. Therefore, the entire photovoltaic array is considered to be in a normal operating state. In case of faulty operation, the detection PNN generates "Class 8" indicating the occurrence of a fault (see Fig. 29), while the diagnosis PNN generates "Class 4, Class 6 and Class 9", which are the corresponding classes for the considered faults diagnosed at the time of their appearance (Fig. 30).

4.3.3.2. PNN classifiers' efficiency under noisy data.

In this last test, the effectiveness of the proposed PNN fault detection and diagnosis is carried out by superimposing the predefined noise levels (Eq. (13)) to the monitored data. The same patterns, described above, have been used in this test. Fault detection and diagnosis results are summarized in Figs. 31 and 32 respectively.

In this particular test, the results shown in Figs. 31 and 32 include some misclassifications mainly due to the exaggerated magnitudes of the superimposed noises. Obviously, such noise levels can lead to highly distorted attributes. But, despite the use of noise signals of a such magnitude, fault detection and diagnosis results based on PNN classifier are very motivating even with the appearance of these misclassifications. In fact, the recorded PNN classifier efficiency is of 82.34% for detection network and 98.19% for diagnosis network.

Table 31. Classification accuracy and Sensitivity results with noisy data

Detection networks tested on the noisy test data			
The classifier type	Classification accuracy %	Sensitivity %	
		Healthy operating case	Faulty operating case
ANN	58.01	61.96	55.07
PNN	82.33	61.96	89.13
Diagnosis networks tested on the noisy data			
The classifier type	Classification accuracy %	Sensitivity %		
		Fault 1	Fault 2	Fault 3
ANN	76.63	69.02	91.85	69.02
PNN	98.19	100	94.57	100

Table 4. Confusions matrixes of ANN networks with noisy data

Detection network type ANN tested on the noisy test data		
The real classes	**The predicted classes**	
	Healthy case	**Faulty case**
Healthy case	123	61
Faulty case	248	304
Diagnosis network type ANN tested on the noisy test data		
The real classes	**The predicted classes**	

	Fault 1	**Fault 2**	**Fault 3**
Fault 1	127	57	0
Fault 2	15	169	0
Fault 3	15	42	127

Table 5 Confusions matrixes of PNN networks for the case of noisy data

Detection network type PNN tested on the noisy test data		
The real classes	**The predicted classes**	
	Healthy case	**Faulty case**
Healthy case	114	70
Faulty case	60	492
Diagnosis network type PNN tested on the noisy test data		
The real classes	**The predicted classes**	

	Fault 1	**Fault 2**	**Fault 3**
Fault 1	184	0	0
Fault 2	10	174	0
Fault 3	0	0	184

Table 6 Classification specificity and Positive Predictivity for ANN networks under noisy data

Detection network type ANN tested on the noisy test data		
classes	**Specificity (%)**	**Positive Predictivity (%)**
Healthy case	55.07	33.15
Faulty case	66.85	83.29
Diagnosis network type ANN tested on the noisy test data		
classes	**Specificity (%)**	**Positive Predictivity (%)**
Fault1	91.85	80.89
Fault2	73.10	63.06
Fault3	100	100

Table 7. Classification specificity and Positive Predictivity for PNN networks under noisy data

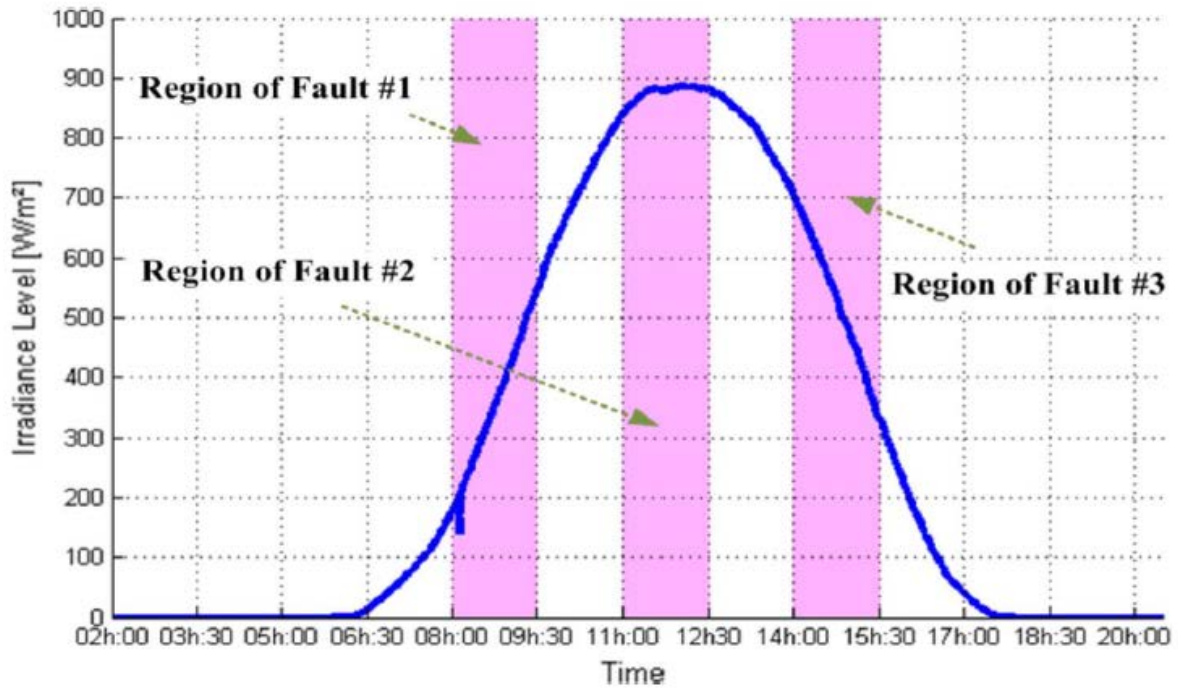| Detection network type PNN tested on the noisy test data		
classes	**Specificity (%)**	**Positive Predictivity (%)**
Healthy case	89.13	65.52
Faulty case	61.96	87.54
Diagnosis network type PNN tested on the noisy test data		
classes	**Specificity (%)**	**Positive Predictivity (%)**
Fault1	97.28	94.85

| Fault2 | 100 | 100 |
| Fault3 | 100 | 100 |

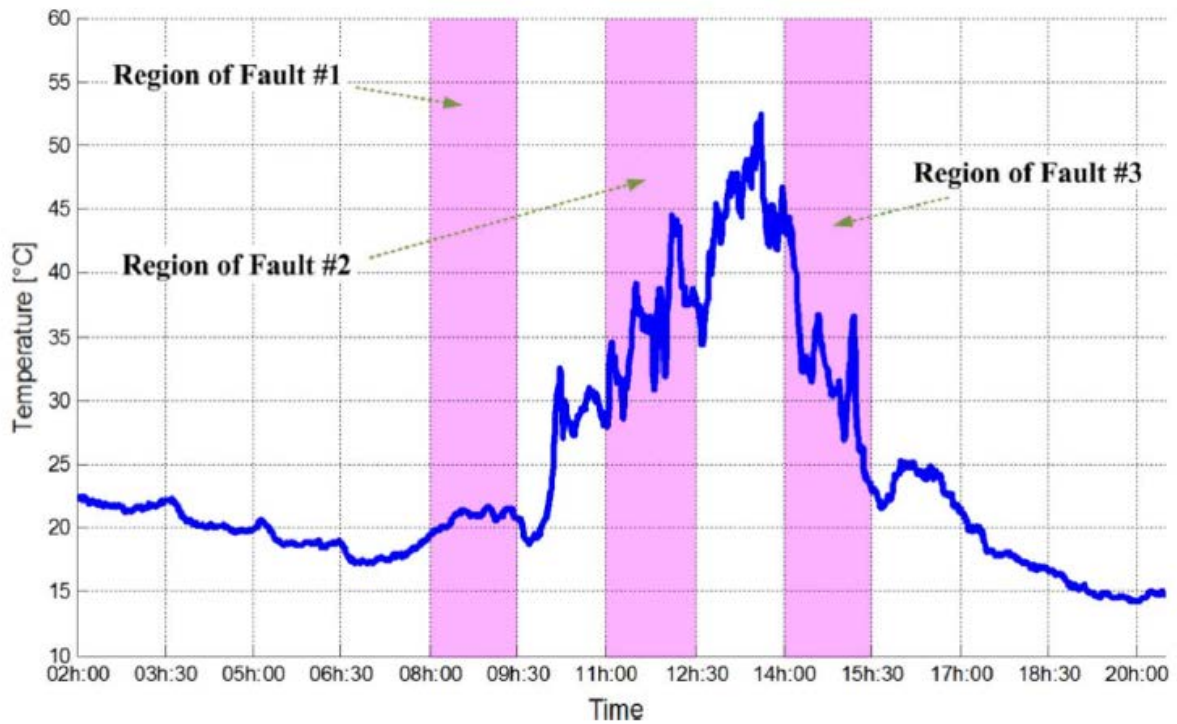Fig. 24. The used daily profile of onsite irradiance level.

Fig. 25. The used daily profile of onsite temperature level.
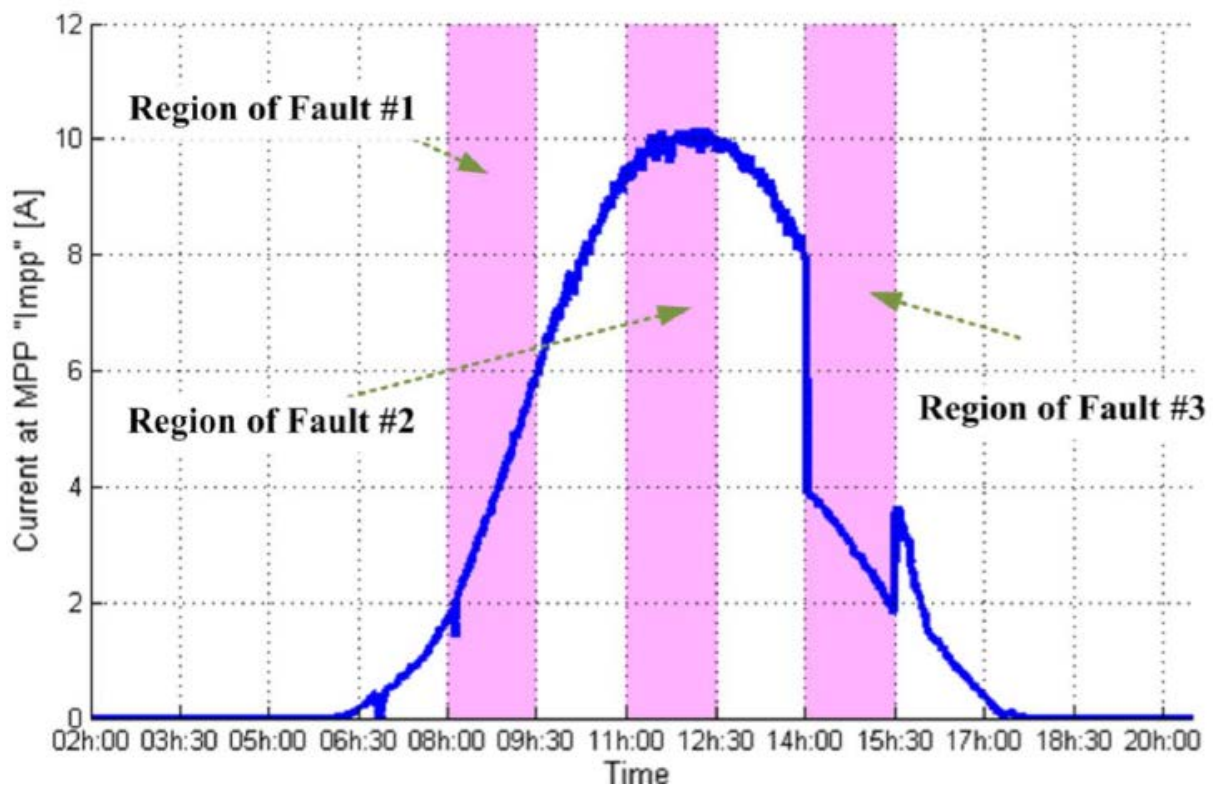
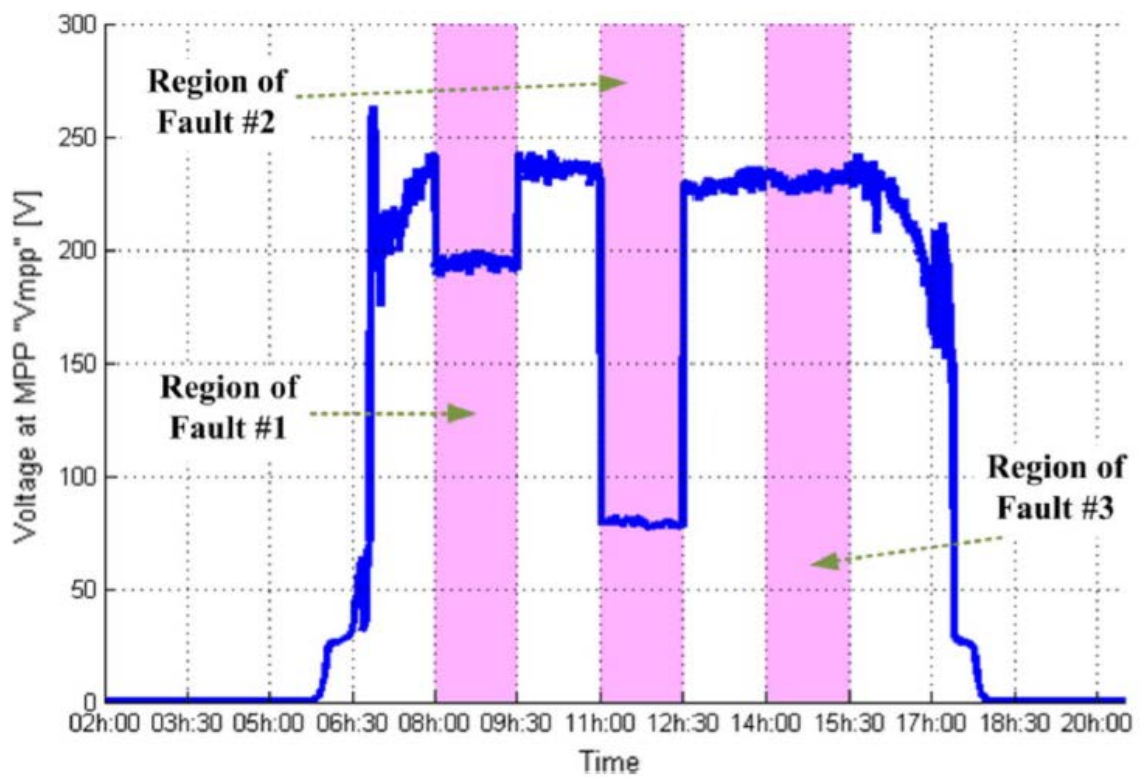Fig. 26. The used daily profile of real measured current at MPP "Impp".

Fig. 27. The used daily profile of real measured voltage at MPP "Vmpp".
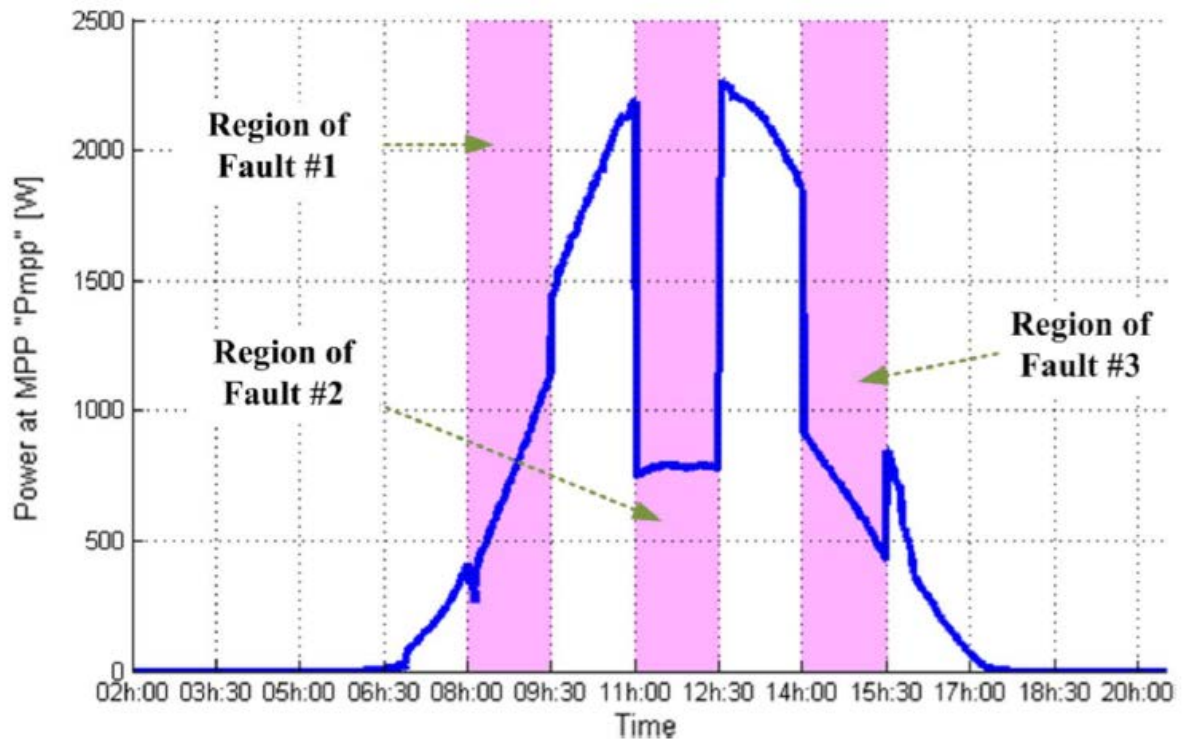
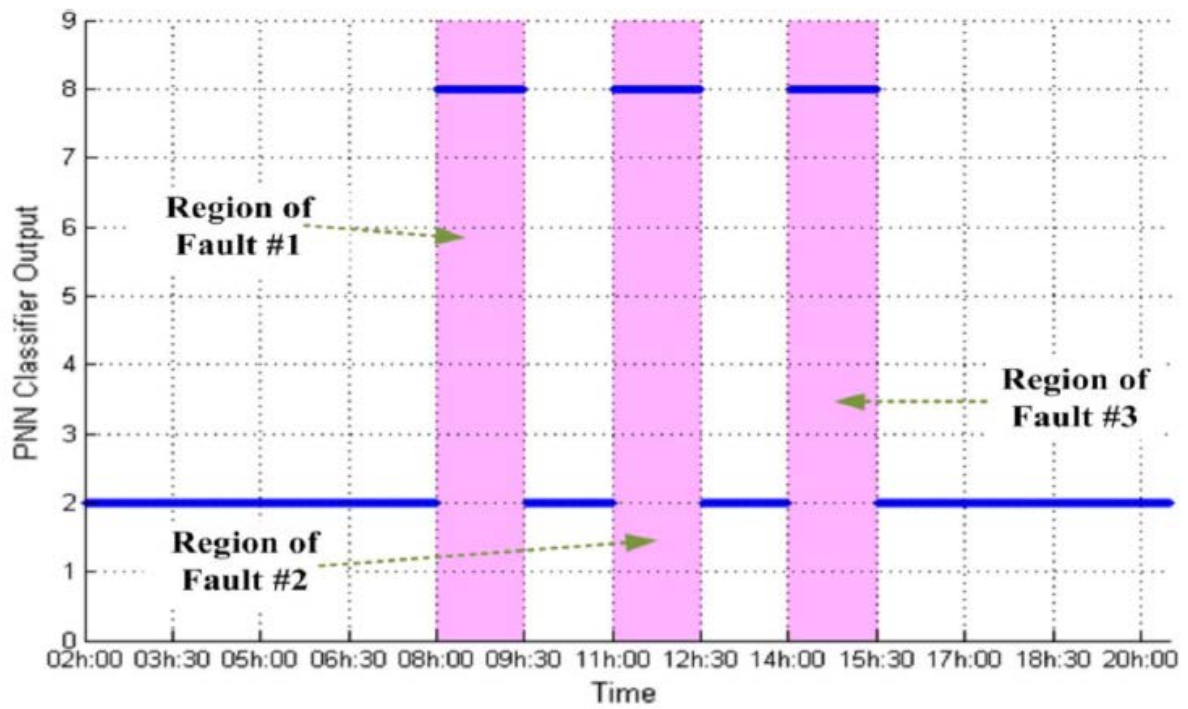Fig. 28. The used daily profile of real measured power at MPP "Pmpp".

Fig. 29. Fault detection results of PNN classifier in the presence of multiple faults and tested under noiseless data.
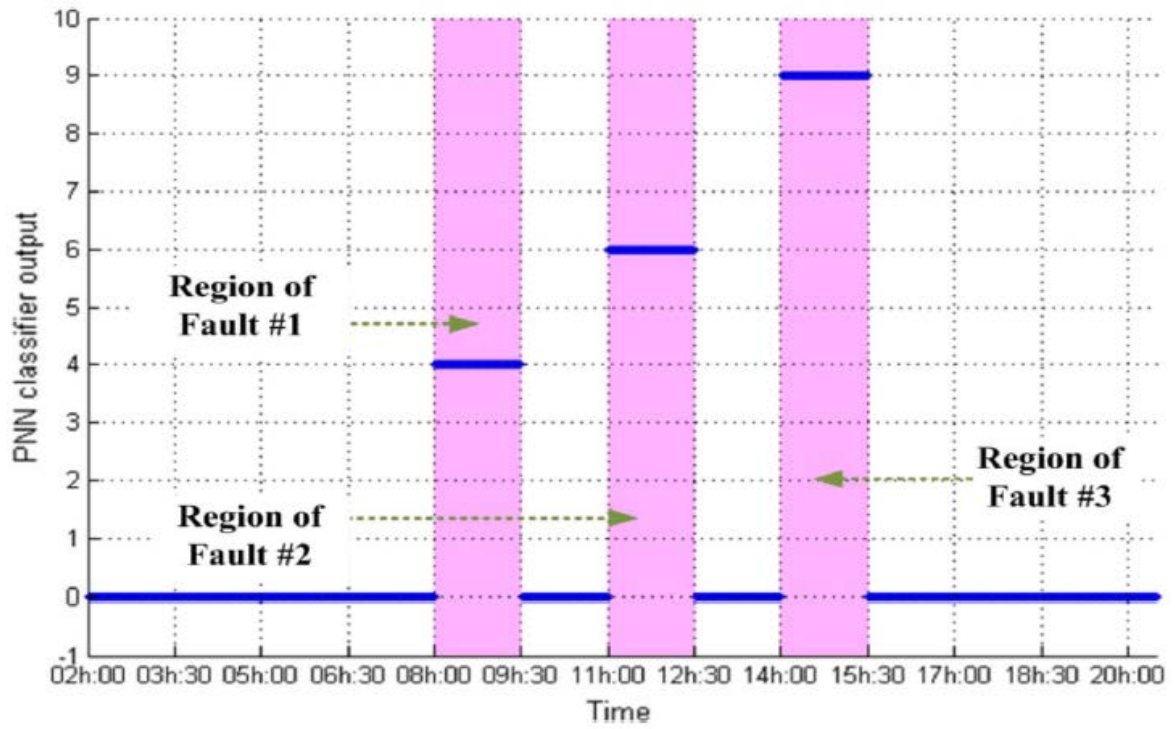

Fig. 30. Fault diagnosis results of PNN classifier in the presence of multiple faults with noiseless data.
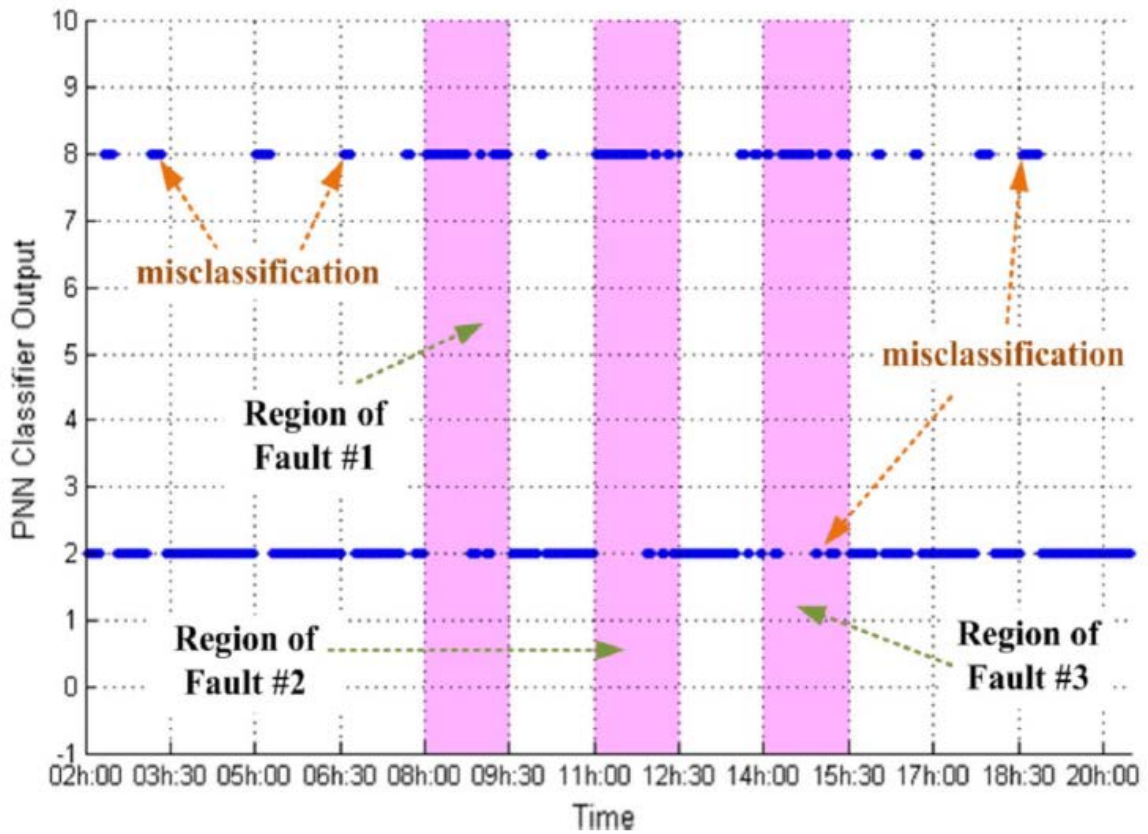
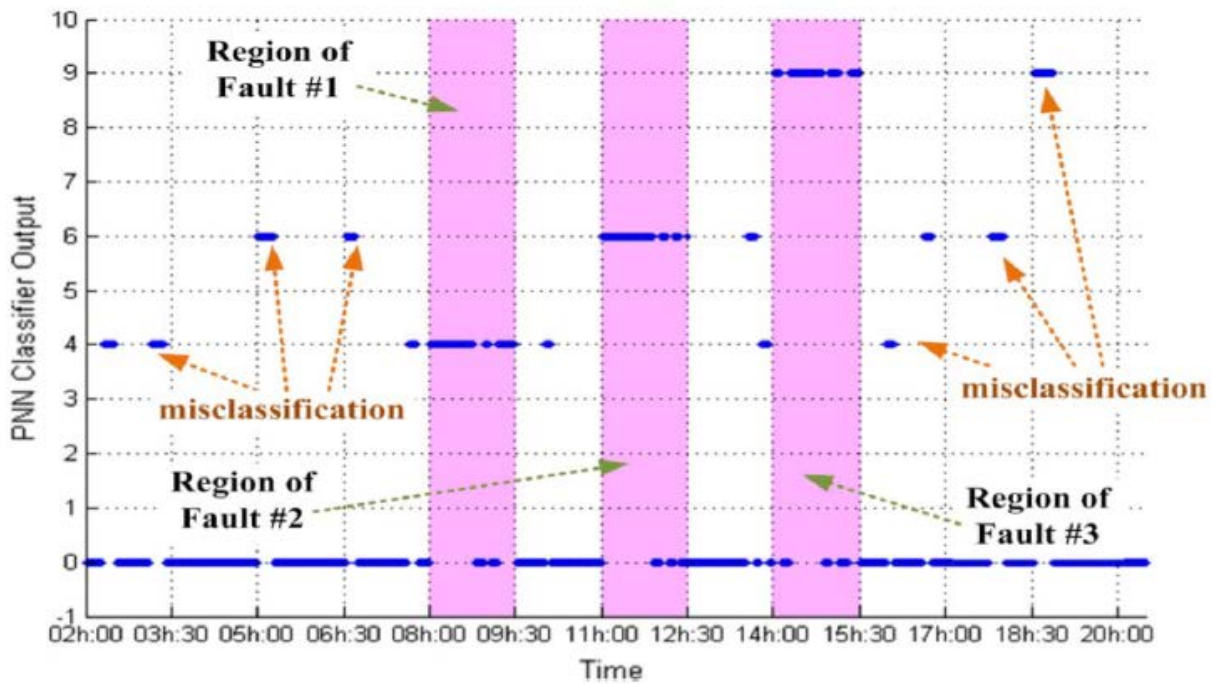Fig. 31. Fault detection results of PNN classifier in the presence of multiple faults with noisy data.

Fig. 32. Fault diagnosis results of PNN classifier in the presence of multiple faults with noisy data.

5. Conclusion

This work has proposed an enhanced machine learning based approach for fault detection and diagnosis of PV systems. The proposed strategy consists of four main stages: PV module parameters extraction based on the best-so-far ABC algorithm; the entire PV system simulation and experimental validation based on a PSIM™/Matlab™ co-simulation; the elaboration of a high-quality database for healthy and faulty operations, and finally, the suggestion of an efficient fault detection and diagnosis procedure based on PNN classifiers.

First of all, the electrical parameters of the used PV module are accurately identified using the best-so-far ABC algorithm. Then, once these parameters are identified, the entire PV system is simulated and experimentally validated using a PSIM™/Matlab™ co-simulation. Finally, the efficiency of the proposed PNN classifier to detect and diagnose failures occurrence has been tested and validated using experimental measurements and simulated data.

Three faulty operating cases of the Algerian PV system have been considered in this work: three modules short-circuited in a PV string, ten modules short-circuited in a PV string and a completely disconnected string from an array. In the other side, the proposed strategy based on PNN classifiers has been tested and compared with conventional feed-forward artificial neural network (ANN) for both noisy and noiseless data.

Finally, this method has been evaluated under real operating conditions using real experimental measurements collected from the tested PV system in order to evaluate its performances and robustness. Results have shown the high efficiency of this method to detect and diagnose DC side anomalies occurrences for both noiseless and noisy data. However, the efficiency of this approach to detect and diagnose shortcircuit faults is guaranteed from at least two short-circuited PV modules. In fact, short-circuit of one PV module cannot be detected using this approach. This is due to its high similarity with healthy system for a large scale PV plant.

References:

[1] Europe S. Global market outlook for solar power 2015–2019. Tech. rep. Bruxelles: European Photovoltaic Industry Association; 2015.

[2] Garoudja, E., Harrou, F., Sun, Y., Kara, K., Chouder, A., & Silvestre, S. (2017). Statistical fault detection in photovoltaic systems. *Solar Energy*, *150*, 485-499.

[3] Madeti, S. R., & Singh, S. N. (2017). Monitoring system for photovoltaic plants: A review. *Renewable and Sustainable Energy Reviews*, *67*, 1180-1207.

[4] Le, P. T., Tsai, H. L., & Lam, T. H. (2016). A wireless visualization monitoring, evaluation system for commercial photovoltaic modules solely in MATLAB/Simulink environment. *Solar Energy*, *140*, 1-11.

[5] Chine, W., Mellit, A., Pavan, A. M., & Kalogirou, S. A. (2014). Fault detection method for grid-connected photovoltaic plants. *Renewable Energy*, *66*, 99-110.

[6] Masson, G., Latour, M., Rekinger, M., Theologitis, I. T., & Papoutsi, M. (2013). Global market outlook for photovoltaics 2013-2017. *European Photovoltaic Industry Association*, 12-32.

[7] Chen, Z., Wu, L., Cheng, S., Lin, P., Wu, Y., & Lin, W. (2017). Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and IV characteristics. *Applied Energy*.

[8] Marion, B., Schaefer, R., Caine, H., & Sanchez, G. (2013). Measured and modeled photovoltaic system energy losses from snow for Colorado and Wisconsin locations. *Solar Energy*, *97*, 112-121.

[9] Potnuru, S. R., Pattabiraman, D., Ganesan, S. I., & Chilakapati, N. (2015). Positioning of PV panels for reduction in line losses and mismatch losses in PV array. *Renewable Energy*, *78*, 264-275.

[10] Reisi, A. R., Moradi, M. H., & Jamasb, S. (2013). Classification and comparison of maximum power point tracking techniques for photovoltaic system: A review. *Renewable and Sustainable Energy Reviews*, *19*, 433-443.

[11] Hariharan R, Chakkarapani M, Ilango GS, Nagamani C. A method to detect photovoltaic array faults and partial shading in PV systems. IEEE J Photovoltaics 2016;6:1278–85.

[12] Daliento S, Chouder A, Guerriero P, Pavan AM, Mellit A, Moeini R, et al. Monitoring, diagnosis, and power forecasting for photovoltaic fields: a review. Int J Photoenergy 2017;2017.

[13] Spagnuolo G, Xiao W, Cecati C. Monitoring, diagnosis, prognosis, and techniques for increasing the lifetime/reliability of photovoltaic systems. IEEE Trans Industr Electron 2015;62:7226–7.

[14] Chouder A, Silvestre S. Automatic supervision and fault detection of PV systems based on power losses analysis. Energy Convers Manage 2010;51:1929–37.

[15] Drews A, De Keizer A, Beyer H, Lorenz E, Betcke J, Van Sark W, et al. Monitoring and remote failure detection of grid-connected PV systems based on satellite observations. Sol Energy 2007;81:548–64.

[16] Karatepe E, Hiyama T. Controlling of artificial neural network for fault diagnosis of photovoltaic array. 2011 16th international conference on intelligent system application to power systems (ISAP). IEEE; 2011. p. 1–6.

[17] Zhao Y, Ball R, Mosesian J, de Palma J-F, Lehman B. Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays. IEEE Trans Power Electron 2015;30:2848–58.

[18] Zhao Y, Yang L, Lehman B, de Palma J-F, Mosesian J, Lyons R. Decision tree-based fault detection and classification in solar photovoltaic arrays. 2012 twenty-seventh annual IEEE applied power electronics conference and exposition (APEC). IEEE; 2012. p. 93–9.

[19] Zhao Y, Lehman B, Ball R, Mosesian J, de Palma J-F. Outlier detection rules for fault detection in solar photovoltaic arrays. 2013 twenty-eighth annual IEEE applied power electronics conference and exposition (APEC). IEEE; 2013. p. 2913–20.

[20] Zhao Y, Balboni F, Arnaud T, Mosesian J, Ball R, Lehman B. Fault experiments in a commercial-scale PV laboratory and fault detection using local outlier factor. 2014 IEEE 40th photovoltaic specialist conference (PVSC). IEEE; 2014. p. 3398–403.

[21] Chine W, Mellit A, Lughi V, Malek A, Sulligoi G, Pavan AM. A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. Renew Energy 2016;90:501–12.

[22] Mekki H, Mellit A, Salhi H. Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules. Simul Model Pract Theory 2016;67:1–13.

[23] Specht DF. Probabilistic neural networks for classification, mapping, or associative memory. IEEE international conference on neural networks. 1988. p. 525–32.

[24] Specht DF. Probabilistic neural networks and the polynomial adaline as complementary techniques for classification. IEEE Trans Neural Networks 1990;1:111–21.

[25] Banharnsakun A, Achalakul T, Sirinaovakul B. The best-so-far selection in artificial bee colony algorithm. Appl Soft Comput 2011;11:2888–901.

[26] Banharnsakun A, Sirinaovakul B, Achalakul T. Job shop scheduling with the best-sofar ABC. Eng Appl Artif Intell 2012;25:583–93.

[27] Banharnsakun A, Tanathong S. Object detection based on template matching through use of best-so-far ABC. Comput Intell Neurosci 2014;2014:7.

[28] Nantapat T, Kaewkamnerdpong B, Achalakul T, Sirinaovakul B. Best-so-far ABC based nanorobot swarm. 2011 international conference on intelligent human-machine systems and cybernetics (IHMSC). IEEE; 2011. p. 226–9.

[29] Garoudja E, Kara K, Chouder A, Silvestre S. Parameters extraction of photovoltaic module for long-term prediction using artifical bee colony optimization. 2015 3rd international conference on control, engineering & information technology (CEIT). IEEE; 2015. p. 1–6.

[30] Garoudja E, Kara K, Chouder A, Silvestre S, Kichou S. Efficient fault detection and diagnosis procedure for photovoltaic systems. 2016 8th international conference on modelling, identification and control (ICMIC). IEEE; 2016. p. 851–6.

[31] Garoudja E, Harrou F, Sun Y, Kara K, Chouder A, Silvestre S. A statistical-based approach for fault detection and diagnosis in a photovoltaic system. 2017 6th international conference on systems and control (ICSC). IEEE; 2017. p. 75–80.

[32] Karaboga D, Basturk B. On the performance of artificial bee colony (ABC) algorithm. Appl Soft Comput 2008;8:687–97.

[33] Boden M. A guide to recurrent neural networks and backpropagation. The Dallas project; 2002.

[34] Dilmac S, Korurek M. ECG heart beat classification method based on modified ABC algorithm. Appl Soft Comput 2015;36:641–55.