Renewable Energy 127 (2018) 548–558

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Multiclass adaptive neuro-fuzzy classifier and feature selection techniques for photovoltaic array fault detection and classification

A. Belaout [a], F. Krim [a], A. Mellit [b, c, *], B. Talbi [a], A. Arabi [d]

[a] Laboratory of Power Electronics and Industrial Control (LEPCI), Faculty of Technology, University of Sétif 1, 19000 Sétif, Algeria
[b] Renewable Energy Laboratory, University of Jijel, Jijel 18000, Algeria
[c] The Abdus Salam International Centre for Theoretical Physics (ICTP), Starada Costiera, 11-34151 Trieste, Italy
[d] Electrical Engineering Department, University of Bouira, 10000 Bouira, Algeria

ARTICLE INFO

Article history:
Received 15 July 2017
Received in revised form
20 March 2018
Accepted 2 May 2018
Available online 3 May 2018

Keywords:
Photovoltaic arrays
Fault detection and classification
Multiclass neuro-fuzzy classifier
Features reduction techniques

ABSTRACT

In this paper, a Multiclass Adaptive Neuro-Fuzzy Classifier (MC-NFC) for fault detection and classification in photovoltaic (PV) array has been developed. Firstly, to show the generalization capability in the automatic faults classification of a PV array (PVA), Fuzzy Logic (FL) classifiers have been built based on experimental datasets. Subsequently, a novel classification system based on Adaptive Neuro-fuzzy Inference System (ANFIS) has been proposed to improve the generalization performance of the FL classifiers. The experiments have been conducted on the basis of collected data from a PVA to classify five kinds of faults. Results showed the advantages of using the fuzzy approach with reduced features over using the entire original chosen features. Then, the designed MC-NFC has been compared with an Artificial Neural Networks (ANN) classifier. Results demonstrated the superiority of the MC-NFC over the ANN-classifier and suggest that further improvements in terms of classification accuracy can be achieved by the proposed classification algorithm; furthermore faults can be also considered for discrimination.

1. Introduction

Photovoltaic systems' reliability is defined in terms of maximizing the electricity production, and minimizing the factors that cause the power losses of the photovoltaic array (PVA), which is related to the implemented strategies for Maximum Power Point Tracking (MPPT) algorithms [1–3] and protection devices avoiding both energy and material losses [4].

At the end of 2016, 303 GW of photovoltaic energy was installed around the world. About 75 GW installed only in the year of 2016 [5], and this is comes from the fact that new solutions have encouraged government to rely more and more on this kind of energy.

Fault diagnosis techniques (FDT) can play an important role to reduce the energy and material losses. Fault diagnosis (FD) in DC-side of a PVA can be classified into three categories, according to the used detection and classification methods, the type of fault to be discriminated, and the chosen features feeding the classifier input:

1) Different types of detection and classification methods were proposed in literature. For example, in Ref. [6] the authors proposed a Neuro-Fuzzy classifier (NFC), by first, constructing an initial Fuzzy Classifier (FC), and then upgrading it with learning algorithms. A database obtained by exhaustive simulation with Matlab/Simulink software. A total of 5790 I-V curves have been stored for features extraction, and further treatment. A FD technique based on artificial neural network (ANN) technique was proposed in Ref. [7]. Two different algorithms have been used; the first one is based on threshold detection and six types of fault, while the second one is based on an ANN. Radial Basis Function (RBF) and Multilayer Perceptron (MLP) architectures have been compared. This later detects and isolates four types of fault [7]. In Ref. [8] a wavelet transform approach has been developed, this method uses also two algorithms, the first one was developed to detect switch open and over harmonic fault using 3-level Multi Level Decomposition (MLD) algorithm, and the second one detects the islanding condition using wavelet coefficients energies. A combination of three FCs, and a sensitivity of the indicators to different type of factors was

[*] Corresponding author. Renewable Energy Laboratory, University of Jijel, Jijel 18000, Algeria.
E-mail addresses: adelmellit2013@gmail.com, amellit@ictp.it (A. Mellit).

https://doi.org/10.1016/j.renene.2018.05.008
0960-1481/© 2018 Elsevier Ltd. All rights reserved.

Nomenclature

ANFIS	adaptive neuro-fuzzy inference system
ANN	artificial neural network
DC	direct current
D_i	output decision of the classifier number i
FC	fuzzy classifier
F_i classifier	classifier of the fault type number i
FIS	fuzzy inference system
I-V	current-voltage
G	irradiance
GW	giga watt
MC	multiclass
MC-NFC	multiclass neuro-fuzzy classifier
MF	membership function
MFN	membership function number
MLD	multi level decomposition
MLP	multilayer perceptron
MPPT	maximum power point tracking
MRE	mean relative error
PCS	Power conditioning system
PV	Photovoltaic
PVA	photovoltaic array
RBF	radial basis function
RMSE	root mean square error
R^2	correlation coefficient
S_i	feature (classifier input) number i
SSE	sum squared error
STD	standard deviation
S_{thi}	threshold for feature (input) S_i
T	temperature

analyzed in Ref. [9], such as solar irradiance and PV module temperature conditions. In Ref. [10] the authors proposed a Graph-Based Semi-Supervised (GBSS) learning algorithm, the model developed has the ability of self-leaning in real-time conditions. Accordingly, it is a low cost model in terms of training. Many other methods can be found in the literature [11–15].

2) The fault types to be detected and classified are varying from one author to other. For example in Ref. [6] the authors proposed a method to detect and then classify three types of faults: increased series resistance, by-pass diode fault, and blocking diode fault. Eight fault types have been detected and classified [7] (PV cells, PV module, PV string and by-pass diode faults). The switch open fault and any over harmonic, and islanding condition were detected and classified in Ref. [8]. These faults occur in a Power Conditioning System (PCS), and it manifests as current-distortion at the output of the PCS. In Ref. [9], partial shading, increased series resistance (ISR) losses and Potential-Induced Degradation (PID) in string PV systems where detected and classified by measurements and extraction of the indicators value of the full PV string I-V characteristic. Ref [10] focuses on two groups of frequently occurring faults in PVAs that cannot be cleared by conventional protection schemes: the line-to-line fault and open-circuit fault. The authors in Ref. [16] have constructed a three state Markov model to represent the state transition relationship of no faults, intermittent faults, and permanent faults for not only PVA, but for all PV components.

3) Selection of classifier input plays an important role to get a high classifier performance for both detection and classification phases. In Ref. [6] the authors have used two features: Maximum Power Point (MPP), and open-circuit voltage (V_{oc}), and when they found it insufficient to discriminate between the considered PV module faults, they add a new ones, namely short-circuit current and filling factor. Five features have been used in Ref. [7]: a reduction in the short circuit current (C_1), A reduction in the open circuit voltage (V_1), a reduction or an increase in the output current (C_2), a reduction or an increase in the output voltage (V_2) and number of peaks in the current-voltage characteristics. The contribution of the previous attributes is analyzed and the most influencing are retained for each type of fault. In Ref. [8] the employed features are: Multi-level Decomposition (MLD) of the wavelet transform, and wavelet coefficients energies are extracted from the grid current and the grid voltage. In Ref. [9] the authors used as inputs classifier: the equivalent thermal voltage (V_{te}), I-V curve flexing factor (IVf), Maximum power point factor (MPPf), Equivalent series resistance (R_{se}) and Fill factor (FF).

For more details, a comprehensive review on fault detection and diagnosis (FDD) methods in photovoltaic systems is published recently in Ref. [17], in which the authors addressed the major PV system failures as well as advantages and limits of FDD methods in terms of feasibility, complexity, cost-effectiveness and generalization capability for large-scale integration.

In the above three categories, the methods have been used for detection and classification of the PV string/array faults, which are mainly based on I-V curves to extract features and run their algorithms. There are many other methods based on visual inspection, and infrared imaging, in which surface soiling, dust accumulation, and hot spot phenomena in PV modules can be detected and removed [18–20].

As reported in Refs. [6–10], fault detection and classification algorithms in a PVA involve the discrimination of many classes at a time. As a solution of this issue, a MC-NFC algorithm is developed, which is mainly based on some strategies that can be adopted in Ref. [21]. One-against-all and one-against-one strategies are the most popular ones. The two strategies are the same in term of classification accuracy. In this paper, the one-against-all strategy is used, which is better than the other in terms of training time, and also for the number of trained classifiers.

An outline of the main contributions of the paper is as follow:

- First, new features have been introduced into the classifier input. Namely, I-V curve area and slopes at different points of the I-V curve.
- Second, unlike in Refs. [6–9], where authors start by some features, and in the case where they found it insufficient to discriminate the total chosen faults, they add new ones. Others like in Ref. [21] used some indicators, and when they found them not meaningful they omit them and seek for new significant parameters. This seems to be done visually by looking to data and using some performance criteria, without using any based-automatic methods for features selection. This paper proposes an inverse method, starting by using many features, and then reducing their number by using variable dimensionality reduction techniques. This alternative solution saves a lot of time for classifiers development.
- Third, in the proposed method some patterns of faults are used, but it can detect all possible real patterns for the concerned

fault, unlike in Ref. [22], where they used some discrete real situation and these patterns could not be exist in real conditions.

- Finally, a Multiclass Adaptive Neuro-Fuzzy Classifier (MC-NFC) has been developed to discriminate between five different types of faults in a PVA.

This study focuses mainly on two objectives. First, feature space dimensionality reduction techniques. Subsequently, a MC-NFC has been designed for detection and classification of five kinds of PVA faults. The rest of the paper is organized as follows: the next section presents the necessary steps in the process of MC-NFC development. Experimental results and discussion are given in Section 3.

2. Multiclass neuro-fuzzy classifierdevelopment

For an efficient and organized manner, six basic steps are necessary in the process of classifier development (See Fig. 1): 1) implement the PVA model; 2) datasets collection covering the most possible scenario for each fault, and the normal operation; 3) features extraction; 4) threshold adjustment for faults detection; 5) classifiers training and testing; 6) classifiers fusion for final decision.

More details about the aforementioned steps will be given in the following paragraphs.

2.1. PV array implementation and data collection

Data sets have been collected by using a real time emulator [23] developed in power electronics and industrial control laboratory (LEPCI) at Sétif-1- University, Algeria (See Fig. 2a and 2b). The advantage of using a real time emulator is that we cannot create faults in a real PVA and we cannot also change the operation conditions. Therefore, in the emulator these constraints can be avoided, and it allows also repeating the same results (I-V curves) and introducing more testing conditions that are not feasible on real PV modules (such as PV cell cracking) (see Fig. 4) (see Fig. 3).

Fig. 2a and b show the schematic representation and the photograph of the used real time emulator, respectively. It consists basically of two parts:

- A software part: Matlab/Simulink and ControlDesk.
- A hardware part: the DS1104 platform which is connected to the PC via PCI slot, and to the APS-1102A programmable DC/AC power source via PLC1104 module.

Current sensor is used to provide a feedback signal to the implemented SPVA model. The control voltage in the external input of the programmable power comes from the controller board. Duty cycle changes regularly from 0 to 1 with a constant slope, then controlling the I-V plotter switch. A resistive load of 5 (ohms)/8 (A) is used.

The PVA model was developed under Matlab software, and then implemented into a DSPACE 1104 control board. The PVA consists of six PV modules connected in series and each module is a connection in series of 36 solar cells (the PV model used is a Bishop model [6,24]). The computation of its parameters is done by the method described in Ref. [25]. The electrical characteristics of the used PV module are given in Table 1.

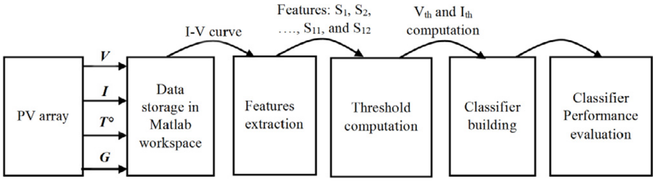

Fig. 1. Process of classifier building for fault detection and classification in a PVA.

As a real time emulator has been used, the data collection period was accelerated by implementing a loop that changed the type of the fault and its severity by considering all possible combinations of solar irradiance and module temperature. The whole dataset was collected within about 6 h divided into two days. In fact, the whole dataset cannot be stored at ones, because of memory limitation in our system.

2.2. Faults in PV array

For the following situations, the I-V curves have been measured by changing solar irradiance and PV module temperature in order to cover all possible real operating conditions, and then the fault is introduced with different discrete values.

✓ PVA without fault (NF): the normal operation condition is in the range of [100 W/m^2, 1000 W/m^2] for the solar irradiance, with step of 100 W/m^2. The module temperature is arranged between [0 °C, 60 °C], with step of 5 °C. The same conditions will be applied in the following situations, but degrees of the fault severity will be added.

✓ Partial shading fault (F_1): Nine different partial shading patterns have been considered. 25%, 50% and 75% of nine PV cell in one PV module, 25%, 50% and 75% of nine PV cell in two PV modules. Finally 25%, 50% and 75% of nine PV cell in three PV modules.

✓ Increased series resistance (F_2): The Rs of one PV module is increased by 1 Ω, 5 Ω, 10 Ω, 15 Ω and 20 Ω.

✓ By-pass diode short-circuited (F_3): One by-pass diode in the whole PVA short-circuited.

✓ By-pass diode impedance (F_4): By-pass diode is assimilated to resistors with different values, 1 Ω, 5 Ω, 10 Ω, 15 Ω and 20 Ω.

✓ PV module short-circuited (F_5): The contribution of one PV module in the energy of PVA was eliminated by making it short-circuit.

By changing the operation conditions (module temperature and solar irradiance), and introducing faults in the plant (single faults), we get more new I-V characteristics. These characteristics are recorded in workspace of Matlab software. Then, we get two matrices (one for the current, and the other one for the voltage). Thus, the real time simulation is finished.

Five datasets are considered, each of them corresponds to one type of fault, and the normal operation dataset (see Table 2). The whole database was split into two equal sets. The first one contains a training data for model construction, and the second one contains a test data for the accuracy estimation of the classifier.

From the previous Table, it can be seen that 2730 I-V curves (faulty PVA), and 130 I-V curves (healthy PVA) have been collected and stored into two distinct matrices, the first one is the current matrix (2860x200), and the second one is the voltage matrix (2860x200). The number 200 comes from the number of points in each I and V vectors.

2.3. Features construction

Feature construction is a process which makes a map from raw data to the classifier input. The aim is to build more efficient features for fault detection and classification task. Voltage and current matrices collected from the PVA, and storage in Matlab workspace for further processing are:

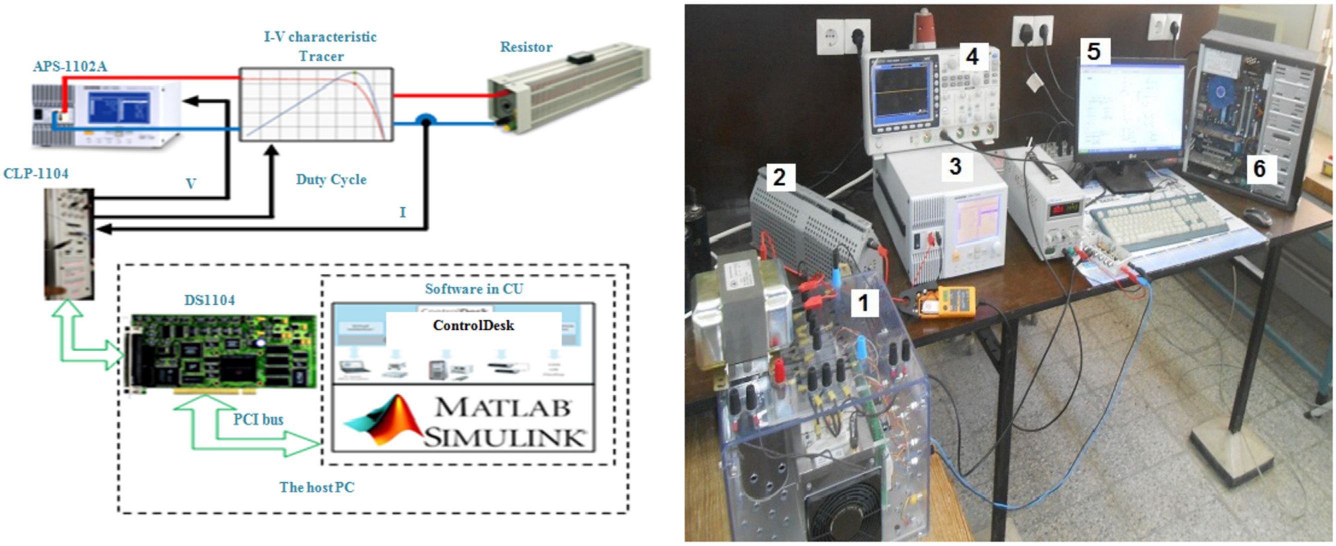

Fig. 2. a. Schematic of the photovoltaic emulator with an APS61102A Programmable Power Source. b. Photograph of the emulator used for data collection: 1) buck converters, 2) load, 3) programmable power source, 4) scope, 5) ControlDesk, 6) ds1104 platform.

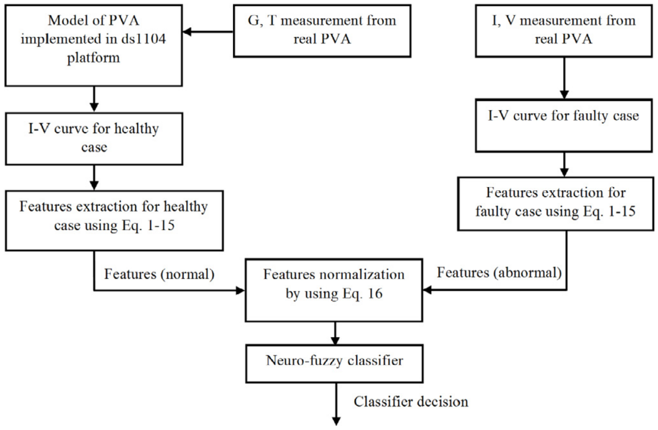
Model of PVA implemented in ds1104 platform
G, T measurement from real PVA
I, V measurement from real PVA
I-V curve for healthy case
I-V curve for faulty case
Features extraction for healthy case using Eq. 1-15
Features extraction for faulty case using Eq. 1-15
Features (normal)
Features normalization by using Eq. 16
Features (abnormal)
Neuro-fuzzy classifier
Classifier decision

Fig. 3. Overview of the proposed method for real-time fault detection and classification phase of unknown samples.

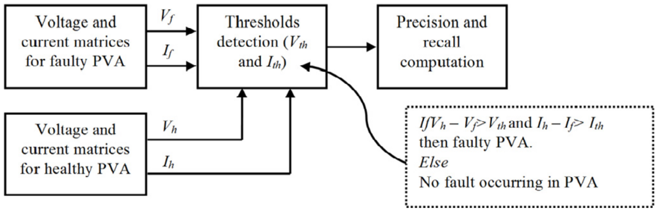
Voltage and current matrices for faulty PVA
V_f
I_f
Thresholds detection (V_{th} and I_{th})
Precision and recall computation
Voltage and current matrices for healthy PVA
V_h
I_h
If $V_h - V_f > V_{th}$ and $I_h - I_f > I_{th}$ then faulty PVA.
Else
No fault occurring in PVA

Fig. 4. Scheme used for precision and recall computation in the detection phase.

Table 1
PV module's electrical characteristics and temperature coefficients (JW-50P).

JW-50P PV module's electrical characteristics	
Number of series cells	36
Open-circuit voltage (Voc)	21.9 V
Voltage at maximum power (Vmp)	17.4 V
Short-circuit current (Isc)	3.13 A
Current at maximum power (Imp)	2.87 A
Maximum power at. STC (Pmax)	50 Wp

Table 2
Datasets for normal and faulty cases of the investigated PVA.

# of Dataset	Fault type	# of samples
0	Normal operation (NF)	130
1	Shading (F_1)	1170
2	Increased series resistance (F_2)	650
3	By-pass diode short-circuited (F_3)	130
4	By-pass diode impedance (F_4)	650
5	PV module short-circuited (F_5)	130

$$V = \begin{bmatrix} V_{11} & \cdots & V_{1n} \\ \vdots & \ddots & \vdots \\ V_{m1} & \cdots & V_{mn} \end{bmatrix} \quad (1)$$

$$I = \begin{bmatrix} I_{11} & \cdots & I_{1n} \\ \vdots & \ddots & \vdots \\ I_{m1} & \cdots & I_{mn} \end{bmatrix} \quad (2)$$

Let us introduce some useful notations for the two matrices. For the *i*-th I-V characteristic, such as *i* runs over 1 to *m*, there are *n* points (in this study, $n = 200$ points) of index *j* for each characteristic (these notations are common for both voltage and current matrices).

From the raw data (I and V matrices) we extract features $S_f (f = 1 \ldots 12)$ for both healthy and faulty PVA using the following formulas:

a) Feature 1: I-V curve area (S_1).

The area under the I-V curve is calculated by the integral:

$$S_1 = \int_0^{V_{oc}} V(I)dI \quad (3)$$

Since the voltage and current vectors (for I-V characteristic) have non-uniform spaced discrete values, one can use the following approximation for numerical implementation of the area of the *i*-th I-V curve:

$$S_1 = \int_0^{V_{oc}} V(I)dI \approx \frac{1}{2}\sum_{j=1}^{n}(I_{j+1} - I_j)(f(I_{j+1}) + f(I_j)) \tag{4}$$

This method is known as the trapezoidal rule.

b) *Feature 2:* short-circuit current *(S_2)*.

This feature can be obtained by the following simple formula:

$$S_2 = I|_{V=0} = I_{sc} \tag{5}$$

From Eq. (5) short-circuit current is the current provided by the PV array where its voltage is equal to zero.

c) *Feature 3:* open-circuit voltage *(S_3)*.

The open-circuit voltage can be calculated as:

$$S_3 = V|_{I=0} = V_{oc} \tag{6}$$

From Eq. (6) open-circuit voltage is the voltage provided by the PV array where its current is equal to zero.

d) *Feature 4:* maximum power point *(S_4)*.

The maximum power of the I-V curve is given as:

$$S_4 = \max(V \times I) \tag{7}$$

e) *Feature 5:* voltage at the maximum power point *(S_5)*. If the index of the voltage V for the *i*-th I-V curve at the maximum power point (MPP) is P, then:

$$S_5 = V|_{V_{ip}} \tag{8}$$

f) *Feature 6:* current at the maximum power point *(S_6)*. V_{ip} is defined in *Feature 5*.

$$S_6 = I|_{V_{ip}} \tag{9}$$

g) *Feature 7*: I-V curve slope at the vicinity of V$_{oc}$ *(S_7)*.

$$S_7 = \left.\frac{dI}{dV}\right|_{V_{oc}} \tag{10}$$

h) *Feature 8:* I-V curve slope at the midpoint between MPP and open-circuit voltage point *(S_8)*. We denote by md1 this first midpoint, so:

$$S_8 = \left.\frac{dI}{dV}\right|_{md1} \tag{11}$$

i) *Feature 9:* I-V curve slope at the MPP *(S_9)*.

$$S_9 = \left.\frac{dI}{dV}\right|_{V_{iP}} \tag{12}$$

j) Feature 10: I–V curve slope at short-circuit current I_{sc} *(S_{10})*

$$S_{10} = \left.\frac{dI}{dV}\right|_{I_{sc}} \tag{13}$$

k) *Feature 11:* I-V curve slope at the midpoint between short-circuit current point and MPP*(S_{11})*. We denote by md2 the second midpoint, so:

$$S_{11} = \left.\frac{dI}{dV}\right|_{md2}. \tag{14}$$

l) Feature 12: filling factor (S12).

$$S_{12} = \frac{(S_5 \times S_6)}{S_2 \times S_3} \tag{15}$$

Features for faulty PVA have been compared to those of normal PVA, and then the results are normalized. In fact, all features for abnormal operation are compared to those for normal operation by using the following formula to get normalized features:

$$feature = \frac{feature(normal) - feature(abnormal)}{feature(normal)} \tag{16}$$

The obtained values for all features are normalized by using Eq. (16), and the final product is a matrix of dimension (2730x12). This latter will be used in the next section for MC-NFC model construction.

As described in Section 2.2, by changing solar irradiance and module temperature in the way to sweep all possible combinations of the operating conditions, we get 130 I-V curves (healthy case). The same conditions have been applied for faulty condition, but degrees of fault severity have been involved. For more clearness, we take the case of Increased series resistance (*F2*): repeating 130 possible combination of couple (G, T) five times (the number of Increased Series Resistance scenarios), the number of normalized features using Eq (16) is 130 × 5 = 650 samples. Thus, the 130 I-V curves for healthy case are used for generating samples by Eq. (16) for all scenarios of different fault types, the raison why we get only 2730 samples at the end.

Once the classifier built, the unknown samples will be normalized using the same Eq. (16). Where feature (normal) is computed from the implemented model in ds1104 platform, and feature (abnormal) is computed from the real PVA, according to the following flowchart:

If unknown samples are presented to the NF classifier, the decision at the classifier output will be “other fault”, indicating that PVA is subject to another fault type that is not considered in this study.

2.4. Threshold detection

To avoid false alarm and the problem of non-detection, a threshold should be chosen for all features, by taking into account measurement errors. For normal operation conditions, the features vector should be inside the hyperbox determined by the following threshold vector:

$$S_{th} = [S_{th1}, S_{th2}, \ldots S_{th12}] \tag{17}$$

Each of the S_{th} vector components corresponds to one feature threshold.

Since the error of measurement is coming mainly from current and voltage sensors, one can compute the error just for these quantities. The remaining threshold components are derived from the current and voltage thresholds.

The IEC 61724-1998 standard [26] tolerates 1% for the current and voltage measurements. The errors for computation of $S_{th2}(I_{th})$ and $S_{th3}(V_{th})$ are the same as the tolerance previously mentioned. Accordingly, S_{th2} (I_{th}) and $S_{th3}(V_{th})$ are used for fault detection.

In the detection phase we focus only on the presence and the absence of the fault in PVA without looking for the type of the fault. In this section, a study of the detection accuracy in terms of precision and recall is conducted (this is a type of binary classification, healthy or faulty PVA).

The following scheme is used to compute precision and recall percentages:

Where V_h, V_f, I_h and V_f are the voltage and the current for the healthy and faulty cases. As reported in Table 2 the number of instances for normal operation is 130 I-V curves, and for the faulty case, the number of instances is 2730 I-V curves. Using these data, the precision and recall percentages are given in the following:

The above scheme for recognizing faulty case in dataset identifies 2703 faulty instances in a dataset containing 2730 faulty instances. Of the 2703 faulty instances identified, 2680 actually are faulty (true positives), while the rest are healthy (false positives). Thus, the detection's precision is 99.15% (2680/2703) while its recall is 98.17% (2680/2730).

2.5. Binary adaptive neuro-fuzzy classifier concept

Fuzzy systems have the ability to handle uncertain and imprecise information, but cannot update and fine tune their parameters automatically. To overcome this drawback, some supervised learning algorithms were applied in Ref. [27], based on training data set. In this study, a well-known Sugeno Fuzzy Inference System (FIS) is used, where its consequent is a constant, this FIS is known as “zero-order Sugeno type” [28].

Before model development, we present a simple architecture that illustrates the procedure of the neuro-fuzzy. Assume we have two inputs, short-circuit current (S_2) and open-circuit voltage (S_3), and one output, increase series resistance. According to the zero-order Sugeno type classifier and in the case where only two rules exist, the output Rs is computed by the summation of the following two functions (f_1 , f_2):

- f_1: is computed by the rule:

$$\text{If } S_2 \text{ is } A_1 \text{ and } S_3 \text{ is } B_1, \text{then} f_1 = p_{11} \times S_2 + p_{12} \times S_3 + r_1 \tag{18}$$

- $f_2$is computed by the rule:

$$\text{If } S_2 \text{ is } A_2 \text{ and } S_3 \text{ is } B_2, \text{then} f_2 = p_{21} \times S_2 + p_{22} \times S_3 + r_2 \tag{19}$$

where p_{ij} and $r_i(i = j = 1, 2)$, are the consequent parameters. Fig. 5 shows a typical ANFIS architecture of such model. Note that in this architecture, squares represent adaptive nodes, whereas circles are fixed nodes.

The word ‘adaptive’ is specific to the NF architecture itself. The adaptive-networks-based fuzzy inference system is a fuzzy inference system implemented in the structure of adaptive network. The output of an adaptive networks depends on the parameter (not features) relevant to the adaptive nodes that is changed to minimize a given error measure using supervised learning procedure [28].

A brief description of the different layers is as follows:

a) Features layer

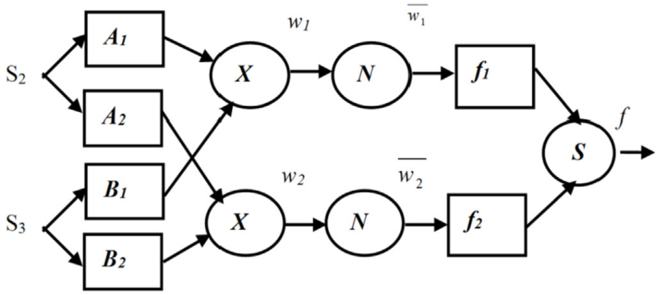

Fig. 5. ANFIS architecture with two inputs, two membership function and one output.

The output of the node i in this layer is calculated by eq.:

$$O_i^1 = \mu_{A1}(S_2)\text{for} \quad i = 1, 2; \tag{20}$$

Or

$$O_i^1 = \mu_{B-2}(S_3)\text{for} \quad i = 1, 2, \tag{21}$$

where, S_2 and S_3 are the crisp features feeding the input of the node $i.A_i$ and B_i are linguistic terms associated with their appropriate membership functions. Membership function for linguistic terms can be any suitable parameterized membership function. Trapezoidal membership function is one of the membership functions that will be used in this study.

$$\mu_{A_i} = \begin{cases} 0, & S_2 \le a_i \\ \frac{S_2 - a_i}{b_i - a_i}, & a_i \le S_2 \le b_i \\ 1, & b_i \le S_2 \le c_i \\ \frac{d_i - S_2}{d_i - c_i}, & c_i \le S_2 \le d_i \\ 0, & d_i \le S_2 \end{cases} \tag{22}$$

$$\mu_{B_i} = \begin{cases} 0, & S_3 \le a_i \\ \frac{S_3 - a_i}{b_i - a_i}, & a_i \le S_3 \le b_i \\ 1, & b_i \le S_3 \le c_i \\ \frac{d_i - S_3}{d_i - c_i}, & c_i \le S_3 \le d_i \\ 0, & d_i \le S_3 \end{cases} \tag{23}$$

where a_i, b_i, c_i and d_i are the parameters to be changed by the training algorithm to deal with training data set. Hence, the trapezoidal function varies consequently.

b) Rules layer

The nodes in this layer provide what is known by firing strengthO_i^2, and it is the product of all outputs coming from layer one. It can be seen that no parameter to be adjusted, so it is a fixed node.

$$O_i^2 = w_i = \mu_{A_i}(S_2)\mu_B(S_3), \quad i = 1, 2 \tag{24}$$

c) Normalization layer

The node i of this layer takes the ratio of the ith rule's firing strength to the sum of all rule's firing strengths. For that reason, outputs of this layer are called normalized firing strength.

$$O_i^3 = \overline{w_i} = \frac{w_i}{\sum_i w_i}, \quad i = 1,\, 2 \tag{25}$$

d) Consequent layer

The output of the node i of this layer is computed by the following node function:

$$O_i^4 = \overline{w_i} f_i \quad i = 1,\, 2 \tag{26}$$

where, $\overline{w_i}$ is a normalized firing strength from the previous layer, and the formula that computes $f_i$is given in Eqs. (17) and (18).

e) Output layer

The single fixed node in this layer computes the overall output by summing all coming signals from the previous layer. Consequently, the process of Defuzzification is achieved by getting a crisp overall output.

$$O_i^5 = ISR = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}, \quad i = 1,\, 2 \tag{27}$$

We note that, the above described ANFIS classifier is just an example. In the following we will use more inputs. More rules will be also generated using learning algorithms. For more explanations on the ANFIS architecture the reader can refer to [28].

2.6. Extension to multiclass classification problem

For a multiclass problem, instead of using one binary classifier we can use a group of them [29]. We can then take their decision and compare it by using “winner-takes-all” rule. The diagram block of the MC-NFC concept is illustrated in Fig. 6.

Each classifier feed this rule by a crisp class label which is assigned to S. The final class label of S is the one that have the biggest crisp value at the outputs of the classifiers pool. If the decision at the output of the classifier i, is d_i, then the label L at the output of the above mentioned rule is given by:

$$L = \max(d_i) \tag{28}$$

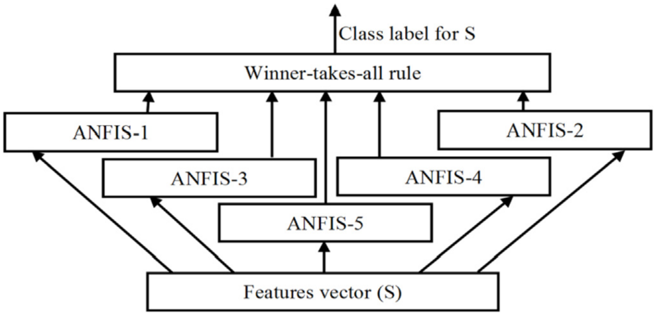

Fig. 6. Diagram block of the MC-NFC concept.

2.7. Classifier performance evaluation criterions

Any constructed classifier should be evaluated at the end of its design, and this allows us to compare it to other type of classifier. Classifier performance will be evaluated by using some statistical criterions: sum squared error, correlation coefficient, mean percent relative error, root mean squared error and standard deviation (See Appendix).

3. Experimental results

The importance of feature dimensionality reduction techniques and the advantage of MC-NFC over traditional ones will be shown. First, in the experiment 1, the proposed method for feature dimensionality reduction has been applied to five classifiers: Partial shading fault classifier (F_1 classifier), increased series resistance classifier (F_2 classifier), By-pass diode short-circuited classifier (F_3 classifier), By-pass diode impedance classifier (F_4 classifier), and PV module short-circuited classifier (F_5 classifier). Then, in the experiment 2, the proposed MC-NFC trained and tested with the reduced entire original feature space, which is defined by a vector of 12 features. Finally, in the experiment 3 the MC-NFC will be compared to an ANN classifier.

3.1. Experiment 1: feature selection for each neuro-fuzzy classifier

First, constructed features have been created without prior information on their effect on the classifier that will be designed. However, what it is known certainly, is that the constructed features will not have the same effect on the classifier decision (output), some of them have the same effect, and some others have no effect on the classifier output at all. These reasons conduct us to reduce the feature space dimensionality. This leads, to avoid redundant information, in the case where some collinear features exist, and also eliminating features that have no effect on output of the decision function.

A standard way to pick a best set of features is via feature space dimensionality reduction techniques. A Matlab program (see Fig. 7) has been developed in order to reduce the classifier input dimensionality. Let us consider n the total number of features, and k the number of selected features, and then the number of trained classifier is given by the following formula (known as combination without repetition):

$$P = \binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{29}$$

The following flowchart explains the algorithm used for inputs

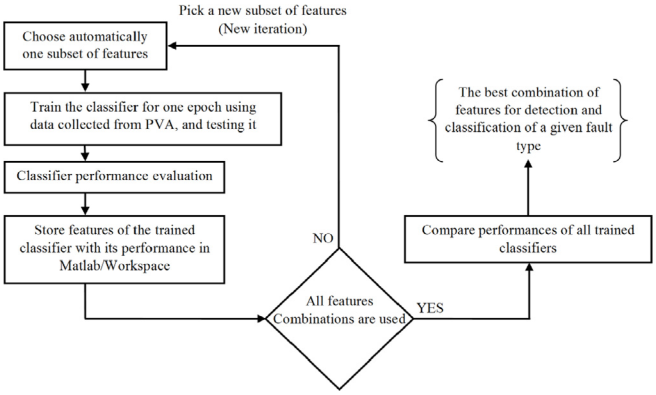
Pick a new subset of features
(New iteration)
Choose automatically one subset of features
Train the classifier for one epoch using data collected from PVA, and testing it
Classifier performance evaluation
Store features of the trained classifier with its performance in Matlab/Workspace
NO
All features Combinations are used
YES
Compare performances of all trained classifiers
The best combination of features for detection and classification of a given fault type

Fig. 7. Overview of the proposed method used for inputs (features) classifier reduction.

(features) classifier reduction.

As indicated in Fig. 7, the algorithm chooses automatically a subset of a 12 features (described in Section 2.3) and train and test the classifier using these latter. Then, the built model will be stored with its performance for further usage. The procedure is repeated until all features combinations are used. At the end, all classifier performances will be compared, and the features will be retained for a given fault type detection and classification.

An ANFIS classifier was built for each combination, and in order to reduce the computation time, the classifier was trained for only one epoch. After that, obtained classifiers must be classified according to their RMSE, and the most relevant combination predicting the output is retained.

As we have five classifiers, the best combination of features is selected for each one of them. In the following, *n* is always equal to 12 (the total number of features), and *k* varies from 1 to 12 (the number of features selected for each combination).

- *F_1 classifier features reduction*

Table 3a demonstrates the result of selecting classifier inputs from one to four, and the best combination of features has been hold. It can be seen clearly that the minimal training (and checking) error are reduced by increasing the number of the selected features. Therefore we will stick to the four-feature classifier for further investigation.

Concerning F_2classifier (See Table 3b) increasing the number of features from 3 to 4 does not minimize significantly the training (checking error), which indicates clearly that the newly added feature does not improve the classification accuracy much. For better generalization, we always prefer a model with fewer inputs. Therefore we will stick to the three-feature classifier for further investigation.

For the same raisons as in F_2 classifier, only two features have been selected for F_3 classifier (See Table 3c).

For the same raisons as in F_2 classifier, only two features are selected for F_4 classifier (See Table 3d).

For the same raisons as in F_2 classifier, only two features are selected for F_5 classifier (See Table 3e).

The results from the above Tables (3a–3e)indicate that the combinations *(S_1, S_3, S_6, S_{12}), (S_4, S_{11}, S_{12}),(S_3, S_4), (S_1, S_4)* and *(S_3, S_4)* are the most influential features with respect to the decision function output the F_1classifier, F_2 classifier, F_3 classifier, F_4 classifier and F_5 classifier, respectively. Consequently, the whole space dimensionality has been reduced from 12 features to only 5 features, namelyS_1, S_2, S_3, S_4, S_{11} and S_{12}.

Fig. 8 shows the detection and classification algorithm basic tasks. First, I-V curve is preprocessed for noise elimination. Second, features extracted from I-V curve by Eqs. (1)–(15) have been normalized by using Eq. (16). Then, switch is positioned on normalized features only if thresholds S_{th2} and S_{th3} are reached, otherwise classifiers will be fed by a zero values for all its inputs components (which means that no fault occurring in the PVA).

Finally classifier output decisions have been compared to decide which one is the biggest for final decision.

Table 3a
Features (inputs) selection for F_1 classifier.
- *F_2 classifier features reduction*

n	12	12	12	12
k	1	2	3	4
RMSE for training	0.6277	0.4183	0.2342	0.1660
RMSE for testing	0.6041	0.4271	0.2046	**0.1740**
Best combination of features	S_3	S_3 S_{11}	S_4 S_8 S_{12}	**S_1 S_3 S_6 S_{12}**

Table 3b
Features (inputs) selection for F_2 classifier.
- *F_3 classifier features reduction*

n	12	12	12	12
k	1	2	3	4
RMSE (Training)	0.7600	0.6298	0.5758	0.5500
RMSE (Testing)	0.7572	0.6319	**0.5539**	0.5200
Best combination of features	S_3	S_4 S_{12}	**S_4 S_{11} S_{12}**	S_2 S_8 S_{11} S_{12}

Table 3c
Features (inputs) selection for F_3 classifier.
- *F_4 classifier features reduction*

n	12	12	12	12
k	1	2	3	4
RMSE (Training)	0.1073	0.0626	0.0356	0.0231
RMSE (Testing)	0.0428	**0.0352**	0.0331	0.0400
Best combination of features	S_3	**S_3 S_4**	S_1 S_3 S_{12}	S_1 S_3 S_8 S_{12}

Table 3d
Features (inputs) selection for F_4 classifier.
- *F_5 classifier features reduction*

n	12	12	12	12
k	1	2	3	4
RMSE (Training)	0.7290	0.5940	0.5521	0.5250
RMSE (Testing)	0.7252	**0.5662**	0.5301	0.5100
Best combination of features	S_4	**S_1 S_4**	S_1 S_4 S_8	S_1 S_2 S_4 S_{11}

Table 3e
Features (inputs) selection for F_5classifier.

n	12	12	12	12
k	1	2	3	4
RMSE (Training)	0.3428	0.0027	0.0012	0.0005
RMSE (Testing)	0.3420	**0.0014**	0.0011	0.0004
Best combination of features	S_1	**S_3 S_4**	S_1 S_3 S_4	S_1 S_3 S_4 S_{12}

Bold number indicate the optimal values of RMSE and the best selected combination.

3.2. Experiment 2: MC-NFC building with reduced feature space

In this experiment, the MC-NFC will be trained and tested with the reduced entire original feature space, which is defined by a vector of 12 features. During the training phase, the ANFIS parameters have been tuned by a hybrid learning algorithm, which consists of a combination of the least-squares method and the back-propagation gradient descent method, and one versus all strategy was applied for all classifiers. The MC-NFC is based on considering the data of the fault to be discriminated as one set, and the remaining data of other faults are forming the other set. Accordingly the problem of multiclass is reduced to a binary classification problem.

An efficient methodology has been adopted to find the best ANFIS model based on the idea of decreasing RMSE and increasing the accuracy of the designed architecture. The main parts of the MC-NFC that will be designed consists on the type of membership functions (Gaussian, triangular, pi-shaped curve, generalized bell-shaped, trapezoidal and differential of two sigmoidal), the number of MFs, output MFs types (constant, linear), the learning process that changes the parameters associated with the MFs and the

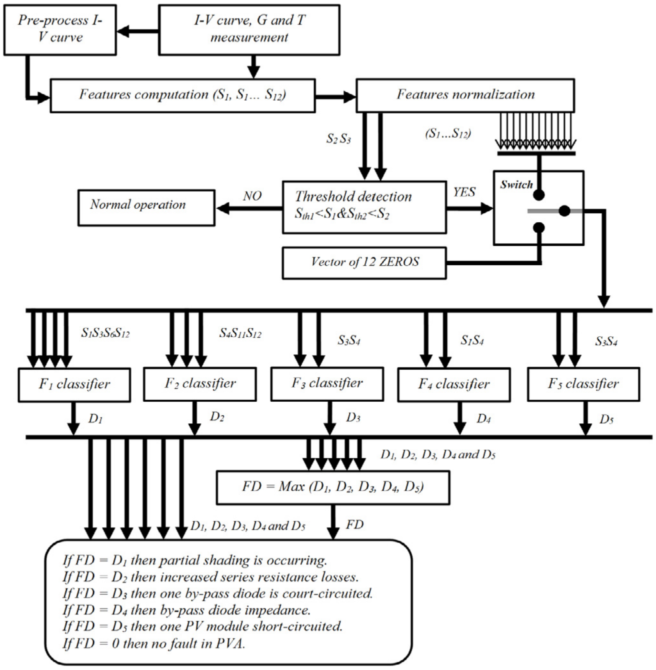
Pre-process I-V curve
I-V curve, G and T measurement
Features computation (S_1, S_1... S_{12})
Features normalization
S_2 S_3
(S_1...S_{12})
Normal operation
NO
Threshold detection S_{th1}<S_1&S_{th2}<S_2
YES
Switch
Vector of 12 ZEROS
$S_1S_3S_6S_{12}$
$S_4S_{11}S_{12}$
S_3S_4
S_1S_4
S_3S_4
F_1 classifier
F_2 classifier
F_3 classifier
F_4 classifier
F_5 classifier
D_1
D_2
D_3
D_4
D_5
D_1, D_2, D_3, D_4 and D_5
FD = Max (D_1, D_2, D_3, D_4, D_5)
D_1, D_2, D_3, D_4 and D_5
FD
If FD = D_1 then partial shading is occurring.
If FD = D_2 then increased series resistance losses.
If FD = D_3 then one by-pass diode is court-circuited.
If FD = D_4 then by-pass diode impedance.
If FD = D_5 then one PV module short-circuited.
If FD = 0 then no fault in PVA.

Fig. 8. Thestructure of the fault detection and classification algorithm based on classifiers decision outputs fusion.

number of epochs to avoid the problem of over fitting.

A grid partitioning method has been applied to generate a rule base relationship between the input and the output of the classifier. The classifier output is a linear combination of its inputs (Sugeno fuzzy inference system).

During the optimization process different built-in membership functions (MFs) types has been involved to choose the most appropriate one for MC-NFC model development. Tables 4a–e demonstrate the errors of MFs during the optimization process. Moreover, the number of MF for each feature has been chosen according to the classifier output performance (See Table 5).

With reference to the above Tables 4 (a–e), it can be summarized that the best membership functions for each classifier is as follows:

About F_1 classifier the best membership function is the differential of two sigmoidal membership function (dsigmf). For F_2 classifier the best membership function is also the differential of two sigmoidal membership function (dsigmf). Concerning F_3 classifier the best membership function is the trapezoidal membership function (trapmf). The best membership function is the generalized bell-shaped membership function (gbellmf) if the case of F_4 classifier. About the F_5 classifier the best membership function is the pi-shaped membership function (pimf).

With respect to Table 5, it can be reported that the best number of membership function for each classifier is as follows:

F_1 classifier: the number of differential of two sigmoidal membership functions is two.
F_2 classifier: the number of differential of two sigmoidal membership functions is four.
F_3 classifier: the number of trapezoidal membership functions is two.
F_4 classifier: the number of generalized bell-shaped membership function is three.

And finally, *F_5*classifierthe number of pi-shaped membership function is two.

Table 4a
Errors of *F_1* classifier membership functions types during the optimization process.

MF-MFN	Membership function type description	MPRE	RMSE	R^2
Gaussmf-2	Gaussian curve	−2.5478	0.16617	0.9724
Trimf-2	Triangular curve	−4.6026	0.18866	0.9644
Pimf-2	Pi-shaped curve	2.0421	0.16169	0.9738
Gbellmf-2	Generalized bell-shaped curve	−1.0257	0.17528	0.9693
Trapmf-2	Trapezoidal curve	2.3275	0.15921	0.9746
Dsigmf-2	difference of two sigmoidal membership functions	1.0529	**0.15022**	0.9774

Bold number indicate the optimal values of RMSE and the best selected combination.

Table 4b
Errors of *F_2* classifier membership functions types during the optimization process.

MF-MFN	Membership function type description	MPRE	RMSE	R^2
Gaussmf-2	Gaussian curve	−5.1359	0.5406	0.7077
Trimf-2	Triangular curve	−8.4089	0.6119	0.6255
Pimf-2	Pi-shaped curve	−3.0729	0.5480	0.6997
Gbellmf-2	Generalized bell-shaped curve	−4.2902	0.5387	0.7098
Trapmf-2	Trapezoidal curve	−1.3245	0.5453	0.7027
Dsigmf-2	difference of two sigmoidal membership functions	−3.2864	**0.5378**	0.7107

Table 4c
Errors of *F_3* classifier membership functions types during the optimization process.

MF-MFN	Membership function type description	MPRE	RMSE	R^2
Gaussmf-2	Gaussian curve	−0.2958	0.0093	0.9999
Trimf-2	Triangular curve	−0.1207	0.0073	0.9999
Pimf-2	Pi-shaped curve	0.0157	0.0020	0.9999
Gbellmf-2	Generalized bell-shaped curve	−0.5259	0.0112	0.9999
Trapmf-2	Trapezoidal curve	0.0054	**0.0018**	0.9999
Dsigmf-2	difference of two sigmoidal membership functions	−0.1370	0.0031	0.9999

Table 4d
Errors of F_4 classifier membership functions types during the optimization process.

MF-MFN	Membership function type description	MPRE	RMSE	R^2
Gaussmf-2	Gaussian curve	−0.6987	0.5272	0.7221
Trimf-2	Triangular curve	−4.6295	0.7077	0.4992
Pimf-2	Pi-shaped curve	−1.5422	0.5308	0.7183
Gbellmf-2	Generalized bell-shaped curve	−2.4022	**0.5256**	0.7237
Trapmf-2	Trapezoidal curve	−1.4081	0.5339	0.7150
Dsigmf-2	difference of two sigmoidal membership functions	−0.7887	0.5278	−0.7887

Bold number indicate the optimal values of RMSE and the best selected combination.

Table 4e
Errors of F_5 classifier membership functions types during the optimization process.

MF-MFN	Membership function type description	MPRE	RMSE	R^2
Gaussmf-2	Gaussian curve	0.0021	0.00016	0.9999
Trimf-2	Triangular curve	−0.3748	0.00670	0.9999
Pimf-2	Pi-shaped curve	−0.0019	**0.00009**	0.9999
Gbellmf-2	Generalized bell-shaped curve	−0.0059	0.00018	0.9999
Trapmf-2	Trapezoidal curve	−0.0001	0.00003	0.9999
Dsigmf-2	difference of two sigmoidal membership functions	−0.0024	0.00010	0.9999

Table 5
Errors of classifiers membership functions number during the optimization process.

	F1-dsigmf		*F2*-dsigmf		*F3*-trapmf		*F4*-gbellmf		*F5*-pimf	
	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
2 MFs	0.235	0.945	0.737	0.456	0.000	1.000	0.526	0.724	0.000	1.000
3 MFs	0.248	0.938	0.626	0.607	0.000	1.000	0.514	0.736	0.000	1.000
4 MFs	0.244	0.940	0.592	0.648	0.000	1.000	0.515	0.734	0.000	1.000
5 MFs	0.30	0.91	0.61	0.63	0.00	1.00	0.52	0.73	0.00	1.00

Bold number indicate the optimal values of RMSE and the best selected combination.

Table 6
Comparison between MC-NFC and ANN-classifier.

	F1		*F2*		*F3*		*F4*		*F5*	
	MC-NFC	ANN	MC-NFC	ANN	MC-NFC	ANN	MC-NFC	ANN	MC-NFC	ANN
RMSE	0.23	0.55	0.59	0.65	0.00	0.31	0.51	0.58	0.00	0.00
R^2	0.95	0.56	0.65	0.66	1.00	0.91	0.74	0.66	1.00	1.00

3.3. Experiment 3: classification with artificial neural network (ANN) classifier

For comparisons purpose, the designed MC-NFC has been compared to ANN classifier (See Table 6). This later is a multilayer feed-forward perceptron (MLP) with one hidden layer. The well-known Levenberg-Marquardt (LM) back-propagation algorithm has been employed.

This experiment shows clearly the superiority of the MC-NFC over traditional ANN-classifier with respect to the reduced feature space dimensionality. Moreover, the proposed MC-NFC can be used to further improve these important results.

4. Conclusions

In this paper, a MC-NFC has been developed for fault detection and classification in photovoltaic arrays. From the conducted experiments, it can be strongly recommended the use of the space dimensionality reduction techniques for classification of PVA faults, due to its capability to speed up the process of classifiers building. This provides to the classifiers a clean manner to select their inputs, high classification accuracy and lower features (inputs) space dimensionality.

The proposed MC-NFC can discriminate between five types of fault occurring in a PVA. Furthermore, the developed algorithm is implemented in a DS1104 platform to show its ability to detect and classify PV array faults in real time applications. Firstly, the original space features was reduced according to their effect of the classifier output. Subsequently, classifiers have been built based on the best combination of the original feature space for each case. Finally, the constructed MC-NFC has been compared to an ANN classifier, and the results show the importance of using the MC-NFC over the traditional one.

Further exploration endeavors the use of other combination of faults, using also other type of classifiers to get more discrimination capacity.

Acknowledgments

Dr. A. Mellit would like to thank the ICTP, Trieste (Italy) for providing the materials and computer facilities for performing part of the present work.

Appendix

- Sum squared error (*SSE*), that is given by the following expression:

$$SSE = \sum_{i=1}^{n} (m_i - p_i)^2$$

- Correlation coefficients (R^2), that is given by the following expression:

$$R^2 = 1 - \frac{SSE}{\sum_{i=1}^{n} p_i^2}$$

- Mean percent relative error (*MPRE*), that is given by the following expression:

$$MPRE = \frac{100\%}{n} \sum_{i=1}^{n} \frac{mi - pi}{p_i}$$

- Root mean squared error (*RMSE*), that is given by the following expression:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (m_i - p_i)^2}{n}}$$

- Standard deviation (*STD*), that is given by the following expression:

$$STD = \sqrt{\frac{\sum_{i=1}^{n} (m_i - p_i)^2}{n - 1}}$$

where m_i is the actual value, p_i is the predicted output of the classifier, and n is the number of the input data.

References

[1] H. Mellit, A. Rezzouk, B. Messai, Medjahed. FPGA-based real time implementation of MPPT-controller for photovoltaic systems, Renew. Energy 36 (2011) 1652–1661.

[2] A. Mellit, S.A. Kalogirou, MPPT-based artificial intelligence techniques for photovoltaic systems and its implementation into field programmable gate array chips: review of current status and future perspectives, Energy 70 (2014) 1–21.

[3] M. Dhimish, V. Holmes, M. Dales, Parallel fault detection algorithm for grid-connected photovoltaic plants, Renew. Energy 113 (2017) 94–111.

[4] Article 690-Solar Photovoltaic Systems, NFPA70, 2016. National Electrical Code.

[5] IEA, Snapshot of Global Photovoltaic Markets, IEA, 2017, pp. T1–T31. PVPS.

[6] A. Belaout, F. Krim, A. Mellit, Neuro-fuzzy classifier for fault detection and classification in photovoltaic module, in: 8th International Conference on Modelling, Identification and Control (ICMIC), Algiers, 2016, pp. 144–149.

[7] W. Chine, A. Mellit, V. Lughi, A. Malek, G. Sulligoi, A. Massi Pavan, A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks, Renew. Energy 90 (2016) 501–512.

[8] S. Kim, On-line fault detection algorithm of a photovoltaic system using wavelet transform, Sol. Energy 126 (2016) 137–145.

[9] S. Spataru, D. Sera, T. Kerekes, R. Teodorescu, Diagnostic method for photovoltaic systems based on light I–V measurements, Sol. Energy 119 (2015) 29–44.

[10] Y. Zhao, R. Ball, J. Mosesian, J.F. de Palma, B. Lehman, Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays, IEEE Trans. Power Electron. 30 (2015) 2848–2858.

[11] M. Dhimish, V. Holmes, M. Dales, Parallel fault detection algorithm for grid-connected photovoltaic plants, Renew. Energy 113 (2017) 94–111.

[12] J. Grichting, M. Goette, Jacomet. Cascaded fuzzy logic based arc fault detection in photovoltaic applications, in: International Conference on Clean Electrical Power (ICCEP), 2015, pp. 178–183.

[13] H. Mekki, A. Mellit, H. Salhi, Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules, Simulat. Model. Pract. Theor. 67 (2016) 1–13.

[14] S.R. Madeti, S.N. Singh, Online fault detection and the economic analysis of grid-connected photovoltaic systems, Energy 134 (2017) 121–135.

[15] L. Zhicong, S. Wu, P. Cheng, Y. Lin, W. Wu, Lin. Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I-V characteristics, Appl. Energy 204 (2017) 912–931.

[16] Y. Cai, Y. Liu, L. Ma, Z. Huang, Liu, A framework for the reliability evaluation of grid-connected photovoltaic systems in the presence of intermittent faults, Energy 93 (2015) 1308–1320.

[17] A. Mellit, M.G. Tina, S.A. Kalogirou, Fault Detection and Diagnosis Methods for Photovoltaic Systems: a Review, Accepted Paper in Renewable and Sustainable Energy Reviews, 2018.

[18] S. Daliento, A. Chouder, P. Guerriero, A. Pavan, A. Mellit, R. Moeini, P. Tricoli, Monitoring, diagnosis, and power forecasting for photovoltaic fields: a review, Int. J. Photoenergy 20 (2017) 17–30.

[19] A. Massi Pavan, A. Mellit, D. De Pieri, S.A. Kalogirou, A comparison between BNN and regression polynomial methods for the evaluation of the effect of soiling in large scale photovoltaic plants, Appl. Energy 108 (2013) 392–401.

[20] J.A. Tsanakas, L.D. Ha, Al Shakarchi, Advanced inspection of photovoltaic installations by aerial triangulation and F. terrestrial georeferencing of thermal/visual imagery, Renew. Energy 102 (2017) 224–233.

[21] W. Hsu, C.J. Lin, A comparison of methods for multiclass support vector machines, IEEE Trans. Neural Network. 13 (2) (2002) 415–425.

[22]] O.,M. Hachana, G. Tina, K.E. Hemsas, PV array fault diagnostic technique for BIPV systems, Energy Build. 126 (2016) 263–274.

[23] A. Belaout, F. Krim, B. Talbi, H. Feroura, A. Laib, S. Bouyahia, A. Arabi, Development of real time emulator for control and diagnosis purpose of Photovoltaic Generator, in: 6th International Conference on Systems and Control (ICSC), Batna, Algeria, 2017, pp. 139–144.

[24] J.W. Bishop, Computer simulation of the effects of electrical mismatches in photovoltaic cell interconnection circuits, Sol. Cell. 25 (1988) 73–89.

[25] D. Sera, R. Teodorescu, P. Rodriguez, PV panel model based on datasheet values, in: IEEE International Symposium on Industrial Electronics, 2007, pp. 2392–2396.

[26] IEC, Photovoltaic system performance monitoring—guidelines for measurement, data exchange and analysis, in: International Standard, IEC, 1998, 61724.

[27] M.F. Møller, A scaled conjugate gradient algorithm for fast supervised learning, Neural Network. 6 (4) (1993) 525–533.

[28] J.S.R. Jang, ANFIS: adaptive-network-based fuzzy inference system, IEEE Trans. Syst. Man, Cybern. 23 (1993) 665–685.

[29] A. Belaout, F. Krim, A. Sahli, A. Mellit, Multi-class neuro-fuzzy classifier for photovoltaic array faults diagnosis, in: In 5th International Conference on Electrical Engineering-boumerdes (ICEE-b), Boumerdes, Algeria, 2017, pp. 1–4.