Maximum power point analysis for partial shading detection and identification in photovoltaic systems

Siwar Fadhel, Demba Diallo, Claude Delpha, Anne Migan-Dubois, I. Bahri, Mohamed Trabelsi, Mohamed Faouzi Mimouni

To cite this version:

Siwar Fadhel, Demba Diallo, Claude Delpha, Anne Migan-Dubois, I. Bahri, et al.. Maximum power point analysis for partial shading detection and identification in photovoltaic systems. Energy Conversion and Management, 2020, 224, pp.113374. 10.1016/j.enconman.2020.113374 . hal-02938015

HAL Id: hal-02938015
https://centralesupelec.hal.science/hal-02938015v1

Submitted on 14 Sep 2022

Maximum Power Point Analysis for Partial Shading Detection and Identification in Photovoltaic Systems

S. Fadhel[1,2,4], D. Diallo[2], C. Delpha[3], A. Migan[2], I. Bahri[2], M. Trabelsi[4], M.F. Mimouni[4]

[1]Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, 4023, Sousse, Tunisie

[2]Université Paris- Saclay, CentraleSupélec, CNRS, Laboratoire Génie Electrique et Electronique de Paris, 91192, Gif-sur-Yvette, France

[3]Université Paris- Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes, 91192, Gif-sur-Yvette, France

[4]Ecole Nationale d'Ingénieurs de Monastir, Laboratoire Automatique, Systèmes Électriques et Environnement, 5035 Monastir, Tunisie

Abstract

Fault diagnosis of photovoltaic (PV) systems is a crucial task to guarantee security, increase productivity, efficiency, and availability. In this regard, numerous diagnosis methods have been developed. Methods requiring the interruption of power production are not adequate for economic reasons. The development of large-scale PV plants and the objective of maintenance cost reduction push toward the emergence of automatic on-line diagnosis methods that use available information. In this study, we propose two data-driven methods for partial shading diagnosis using only the maximum power point's information. It does not require the interruption of production, nor does it require any additional equipment to obtain the I(V) curve. The analyses are conducted with principal component analysis (PCA) and linear discriminant analysis (LDA) to detect and classify the faults. The experimental dataset is collected from a 250 Wp PV module under four states of health (healthy, and three severities of partial shading) for several meteorological conditions. The classification results have a 100% success rate, and are robust to the variations of temperature and irradiance.

Keywords: Photovoltaic system, Fault diagnosis, Partial shading, Maximum Power Point (MPP), Principal Component Analysis, Linear Discriminant Analysis

1. Introduction

Although photovoltaic (PV) cells/modules are the most reliable components in a PV installation [1] and considered as systems with low failure rates [2] requiring low maintenance [3], it is crucial to implement effective monitoring systems for PV modules to diagnose their state of health. Indeed, many PV faults can cause significant power losses, may lead to accelerated system ageing, reduced efficiency, and even safety hazards.

Due to uncertainties related to PV cell technology, manufacturing process and climatic conditions, correlating PV fault type to power loss is tricky. It is still an on-going area of research. In particular, the impact of climatic conditions is evaluated in [4], [5] by studying a large number of PV modules exposed to different climates. The annual degradation rate is proposed by numerous studies to quantify the impact of single or multiple faults. For example, it has been reported that crystalline silicon (c-Si) PV modules have degraded at an annual rate of approximately 7% due to discoloration and weakened solder bonds [6]. Another study reported that because of discoloration, the power loss could reach 10% per year in PV modules with concentrators [5]. According to [4], PV installations deployed after the 2000s are more susceptible to the occurrence of defects. However, the percentage of degradation would be significantly lower than in previous installations. In particular, for new c-Si PV modules, the biggest concern seems to be hot spots, followed by potential induced degradation (PID). Hot spots have an impact on PV module performance and could contribute to premature ageing if they are frequent [7]. It has also been reported that partial shading (PS) can lead to several early degradations, and is a significant cause of the appearance of hot spots. Therefore, we are interested in partial shading (PS) diagnosis.

Several PV faults have a visual signature such as delamination, burn, or discoloration marks. So, they can be simply detected through visual inspection. This technique is clearly irrelevant for large PV installations. Moreover maintenance policies are evolving towards Condition Based Maintenance

(CBM) to prevent failures and reduce maintenance costs. CBM requires continuous monitoring and automatic fault diagnosis methods.

There are several methods for diagnosing PV faults, reviewed in the literature [8] to [11]. They can be classified into two main families according to the measures used:

- The first group uses radiation-based techniques: Electroluminescence (EL) [12], [13], and Infrared Thermography (IRT) under steady-state conditions [13], [14]. The main advantage of these techniques is their ability to detect and locate the faults in real-time. EL is generally performed to detect micro-cracks in PV cells and other faults within the cell material (contact degradation…); however, it requires the system interruption. IRT can be performed at module level, allowing a fast and effective inspection to detect different types of faults. This technique can be tested without any operational interruption, but it requires mobile platforms to diagnose many PV modules. For this reason, researches are ongoing to develop advanced IRT techniques such as drone-mounted thermography [15]. Therefore, IRT could be more suitable than the EL technique, especially for big PV plants.
- The second group uses electrical measures. It is composed of reflectometry techniques, I(V) curve analysis, and maximum power point (MPP) coordinates analysis.

Reflectometry techniques are mainly used to detect catastrophic faults (ground faults and arc faults) and locate open-circuit faults in the PV strings. These methods are effective. Nevertheless, they are costly since they require a specific external signal function generator to stimulate the PV system [16]–[20].

I(V) curves can also be used for fault diagnosis. However, not only do several defects have very similar signatures, but the I(V) curve is also very sensitive to environmental conditions. Therefore diagnosis methods based on the analysis of the entire I(V) curve's points need appropriate preprocessing. This difficulty may be one of the reasons that justify the limited number of publications using the full I(V) curve for fault diagnosis. From the literature review, we have only found two studies dealing with the full I(V) curve analysis for diagnosis purposes [7], [23]. In [7], the analysis of the first derivative of the ratio between the standard error and the PV voltage is used for the same fault, the standard error results from the comparison between faulty and healthy I(V) data. Authors in [23] have proposed the evaluation of the first and the second derivatives of the curve for shading detection. An I(V) plotter has been used in this study to sweep the I(V) curve composed of 265 points. We have proposed in two previous works a new method based on the full I(V) curve analysis to detect and classify shading faults by using adequate variables [21], [22]. The experimental I(V) characteristics are obtained thanks to a programmable voltage source, which provides the points from the short-circuit current to the open-circuit voltage.

The other diagnosis techniques use the actual I(V) curve to extract several health indicators from the shape of the curve. The fault diagnosis decision is made upon analysis of the residuals computed as differences between measured indicators and estimated ones identified from the healthy I(V) characteristics. This latter is generally obtained by considering the usual electrical equivalent circuit model (single-diode or double-diode) [24]. At first, an optimization algorithm is applied to extract the five optimal parameters, as proposed in [25]–[29]. Then, the non-linear equation I = f(V) is solved iteratively to reproduce the healthy behavior of the PV system.

In [30], the authors have defined five indicators (fill factor, equivalent series resistance…) calculated from the shape of the outdoor measured on-line I(V) curve. Fault occurrence is detected by comparing the actual values with their initial values measured during the commissioning phase of the diagnosis system. Based on the one-diode model, authors in [31] have defined two reference indicators to estimate the increment of the series resistance from the measures of the short-circuit current I_{SC}, PV system temperature T_{PV}, maximum power point's voltage and current. The experimental indicators analysis is performed with a variable load used to extract the full I(V) curve.

These methods suffer from several difficulties. First, they introduce additional time to sweep the full I(V) characteristic, and the duration and the quality of acquisition depend on the used instrument. Second, these approaches are confronted with the case of large-scale photovoltaic plants with limited measurements. Furthermore, it is not practical to disconnect PV modules or strings to extract their I(V) curves as it perturbs the system operation and leads to production losses. For instance, the power loss due to the I (V) acquisition in [32] resulted in a power efficiency drop of 0.16%.

For this reason, it is not suitable for continuous monitoring. There is therefore, an interest in using methods that do not require the interruption of production and that use only available data such as voltage and current at the point of maximum power (MPP) commonly available in PV systems. This approach is also cost-effective, as it does not require any additional measurement equipment.

The fault detection strategy developed in [33] is based on monitoring only the maximum power of the PV system by analyzing the sample entropy. In [34], authors have proposed two diagnosis attributes, which are the voltage ratio (VR) and power ratio (PR) computed from the division of the theoretical and the measured values of the maximum power point's voltage and power, respectively. More recently, to distinguish healthy from faulty states, three diagnosis indicators calculated from the comparison between predicted and measured current, voltage, and power at MPP have been proposed in [26]. The same electrical measurements are exploited in [35], [36], to define three residuals as fault features. The data in these two studies are recorded from the same PV system whose behavior in the healthy state is estimated based on the one-diode model. Residuals are analyzed with the one-class support vector machine (1SVM) technique in [35], while the exponentially weighted moving average (EWMA) approach is applied to the same residuals in [36], to detect and identify the faults. All the previous analysis methods using the maximum power point's coordinates rely on physics-based models. These methods are powerful in detecting many types of PV faults and can be applied to a wide range of PV plants. However, they rely on an analytical model to reproduce the nominal PV characteristics, and an optimization process is implemented for parameter extraction. Subsequently, MPP's coordinates in a healthy state are estimated under real climatic conditions. When dealing with large PV plants, the modeling stage can be quite challenging to reproduce the MPP's coordinates with high accuracy. Besides, due to the sensitivity of PV systems to weather conditions, optimal thresholds must be selected for effective and robust detection. Moreover, physics-based models need to be updated to take into account the natural ageing and degradation of solar cells. Unfortunately, these mechanisms are not yet well known in order to develop reliable models. Data-driven methods based on historical data could be a valuable alternative for PV fault diagnosis.

To the best of our knowledge, fault detection based on data-driven approaches using only real data at Maximum Power Point (MPP) data has not been well addressed; only one publication has been found in the literature [3]. Authors have proposed graph-based semi-supervised learning for fault diagnosis. The method is able to detect and identify line-to-line and open-circuit faults efficiently, but it is not suitable for partial shading detection.

It should be remembered that partial shading (PS) is responsible for several early degradations, and is an essential cause of the appearance of hot spots. That is why, in what follows, we focus on the diagnosis of partial shading (PS) using a methodology based on maximum power point data.

The methodology is based on four steps:

- Modelling step in which the data is collected from an experimental PV module of 250 Wp,
- Preprocessing step in which the variables (voltage, current, and power at the maximum power point) are selected, centered, reduced and processed through the logarithmic function,
- Features extraction: in this step two statistical techniques, principal component analysis (PCA) and linear discriminant analysis (LDA) are applied.

- Features analyses: in this step, PCA and LDA are used for partial shading faults detection and classification. LDA is also used to determine discriminative laws for data separation.

We will evaluate the robustness of the method to environmental conditions: variable temperature and irradiation.

The outline of the paper is as follows: section 2 describes how the database is collected and extended. The detailed faults diagnosis methodology based on maximum power point (MPP) coordinates are presented in section 3. A case study of an isolated PV module is presented to validate the proposed methods. Section 4 discusses the application to large-scale PV plants and compares its practical implementation with full I (V)-based diagnosis technique. Finally, section 5 concludes the paper.

2. Description of the dataset

2.1.Data acquisition within a 15-minute time window

The experiments were carried out on March 21, 2019, using the experimental setup displayed in Fig. 1. The main parameters of this module at standard test conditions (STC) are as follows [22]: the maximum power P_{mpp} = 250 Wp, the maximum power current I_{mpp} = 8.21 A, the maximum power voltage V_{mpp} = 30.52 V, the short-circuit current I_{sc} = 8.64 A, the open-circuit voltage V_{oc} =37.67 V. The temperature of the PV module T_{PV} is measured with a 4-wire Pt100 probe bonded on its back face.

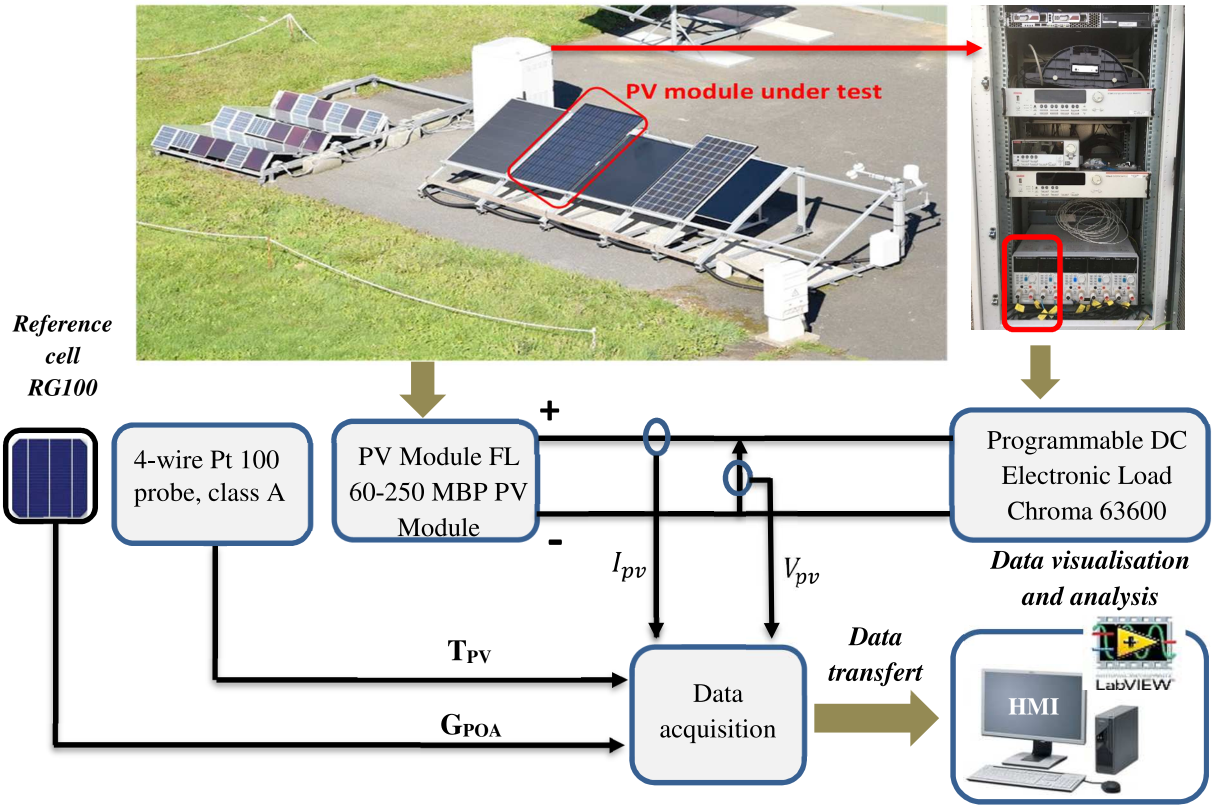


Fig. 1. Installed PV panel platform

A reference cell (SOLEMS RG100) is used to measure G_{POA} the plane of array irradiance.

The acquisition, display, and saving of each I (V) characteristic take one minute. Therefore, during 15 minutes, 15 experimental I(V) curves for each health status (healthy and the different shading profiles) are recorded for several levels of temperature T_{PV} and irradiance G_{POA}. Each curve is composed of 101 points. As an example, for the environmental conditions T_{PV} and G_{POA} presented in Fig. 2, the characteristics I(V) and P(V) under healthy status are displayed in Fig. 3. We can observe that due to the temperature and irradiance fluctuations $((\Delta G_{POA})max = 25\%$ and $(\Delta T_{PV})max = 13\%)$, the maximum power fluctuates from 149 Wp to 190 Wp.

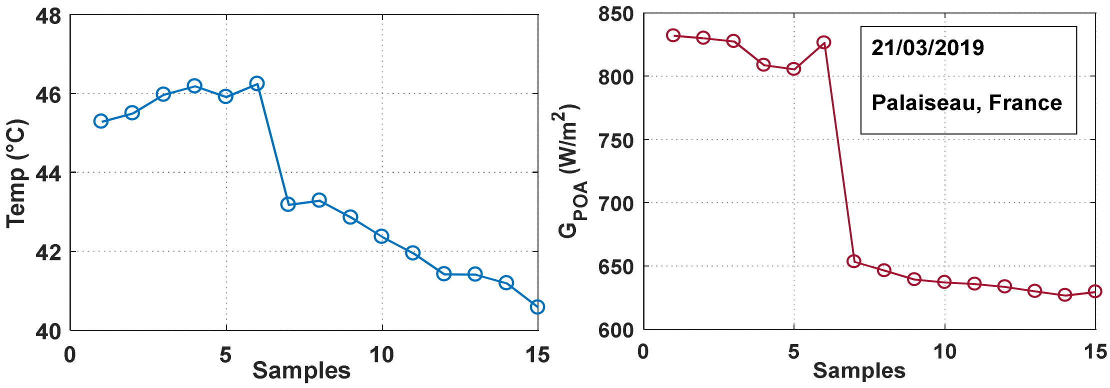


Fig. 2. Module temperature and irradiance variations for a healthy PV module

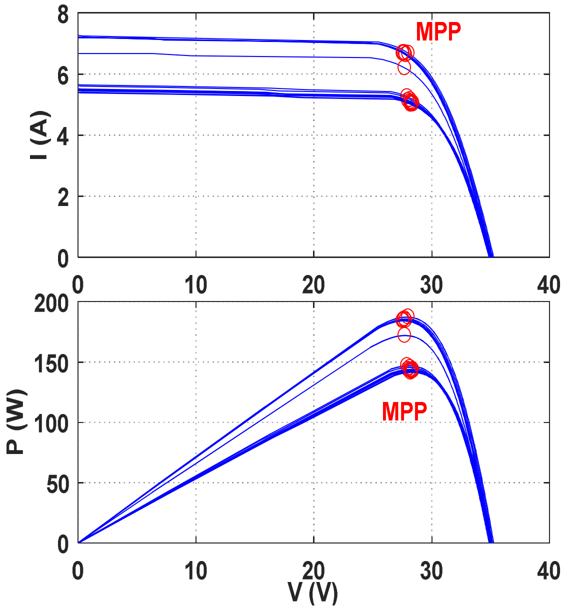
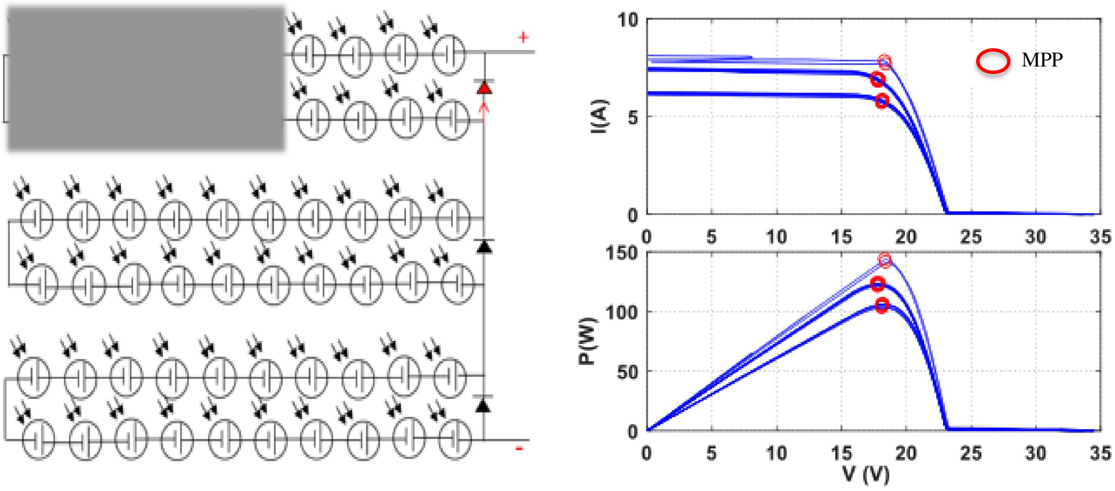


Fig. 3. Measured I(V) and P(V) curves for a healthy PV module

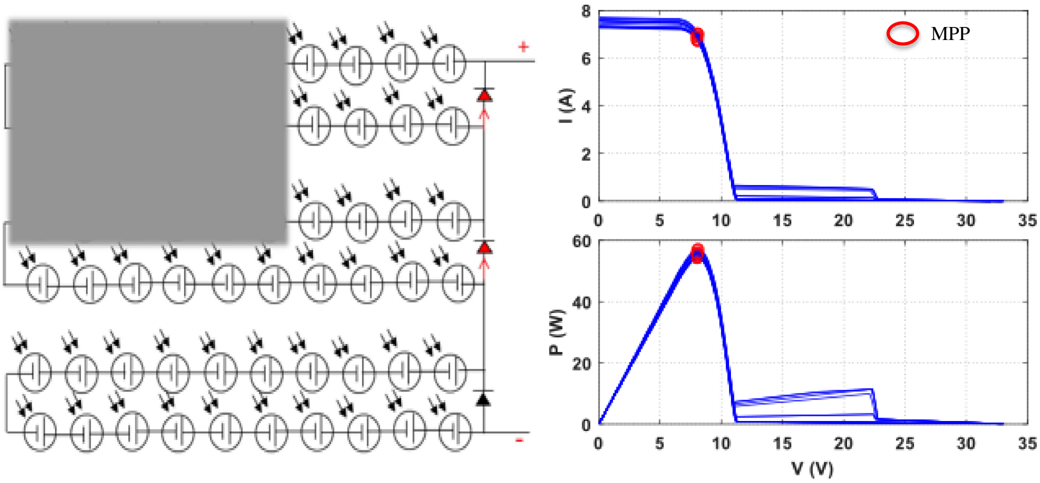
The shading profiles as in [22], and the corresponding experimental I(V) and P(V) curves are displayed in Fig. 4. One can notice the maximum power drop caused by the bypass diodes activation. The solar irradiance and the module temperature measured during the partial shading experiments are plotted in Fig. 5. The maximum relative variations of irradiation and temperature are summarized in Table 1.

Table 1. Maximum relative variations of irradiance and temperature

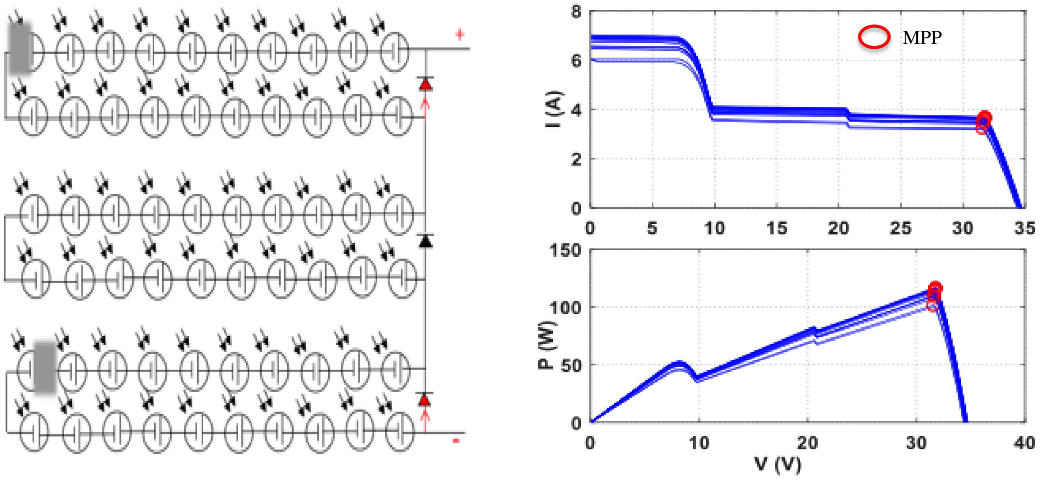
Partial Shading Condition	$\Delta(G_{POA})$max (%)	$\Delta(T_{PV})$max (%)
1	36	18.5
2	6	32
3	16	4.5



(a) Partial shading 1

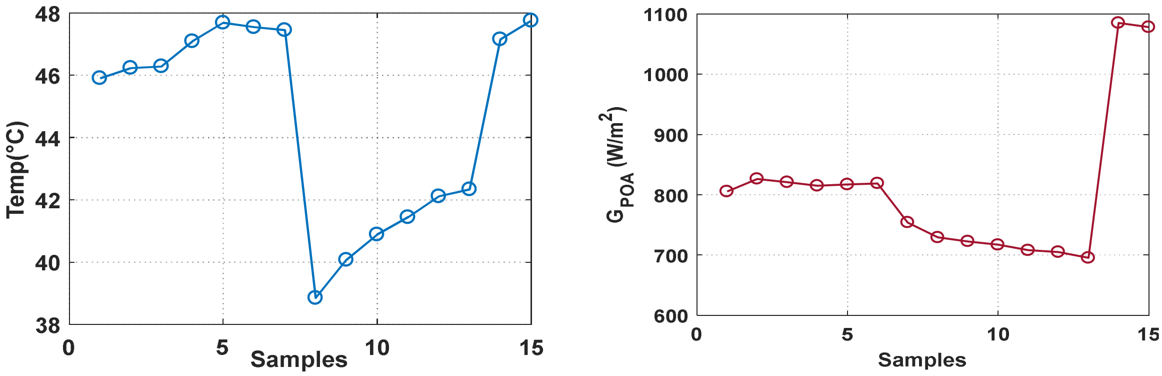


(b) Partial shading 2

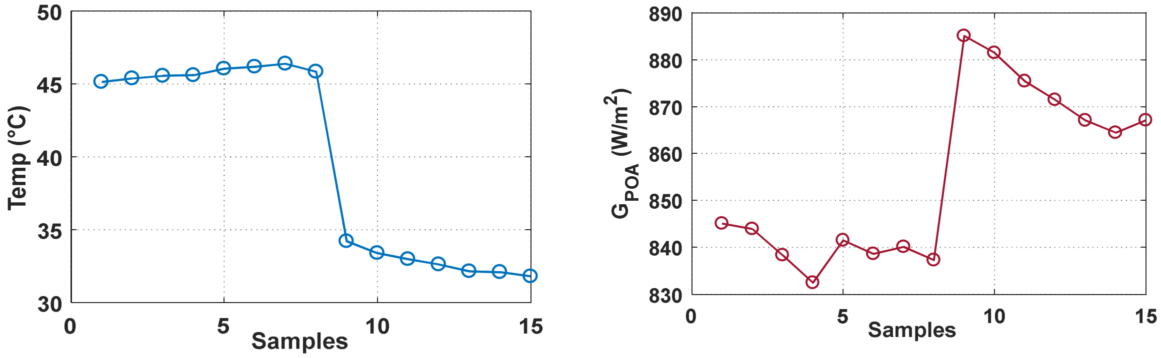


(c) Partial shading 3

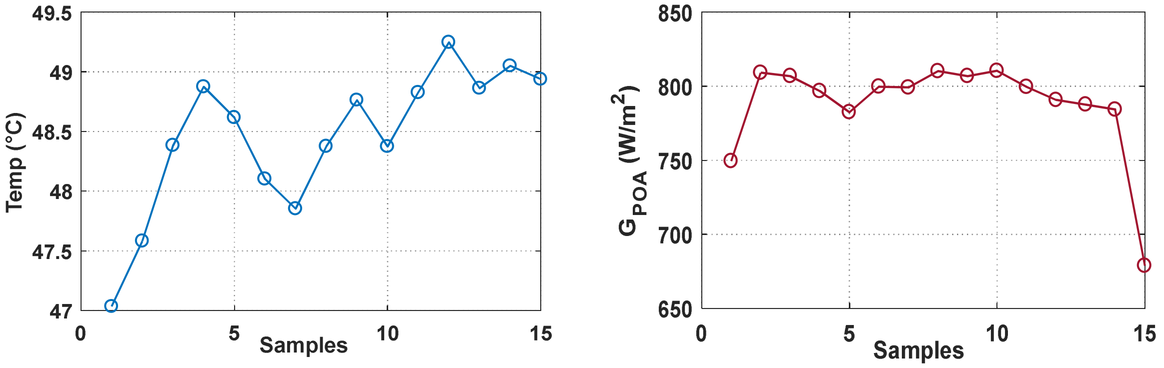
Fig. 4. Partial Shading profiles and measured I (V) and P (V) curves



(a) Partial Shading 1



(b) Partial Shading 2



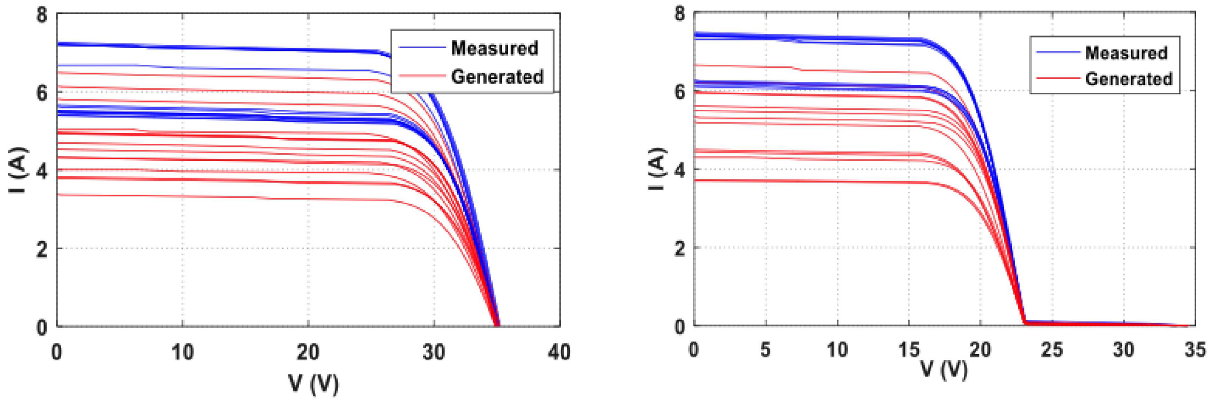
(c) Partial Shading 3

Fig. 5. Module temperature and irradiance variations for the different partial shading conditions

2.2. Extension of the database

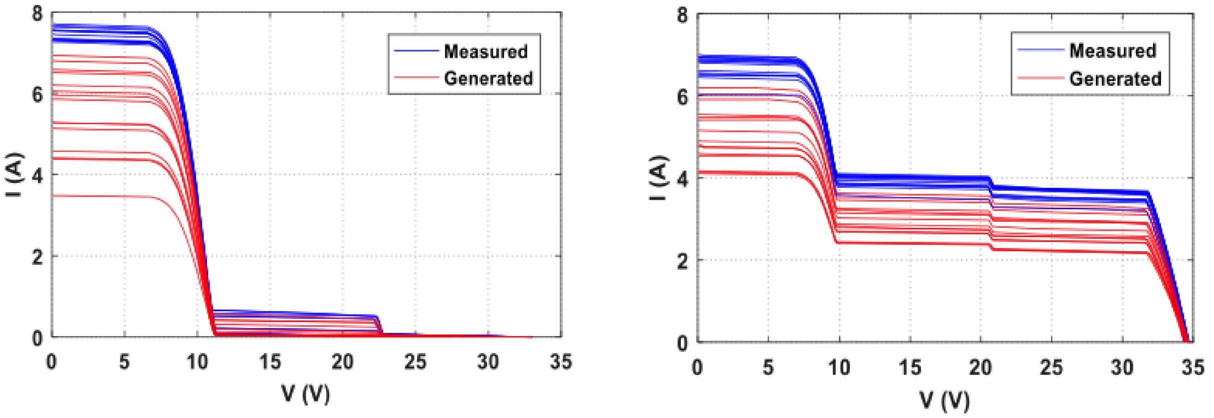
To increase the number of points in the database, we have generated new I(V) curves from the experimental ones by varying the solar irradiance. In the following, we have doubled the number of points; from the original 15 maximum power points for each health status, we have now 30 points. Therefore, for the four modes (one healthy state and three partial shading ones), we have 30*4 MPP observations. Experimental and generated I(V) curves are plotted respectively with blue and red colors in Fig. 6.

It must be noticed that MPP coordinates are almost always available in PV plants under operation, thanks to the global maximum power point tracking (GMPPT) algorithm [37, 39].



Healthy

Partial Shading 1



Partial Shading 2

Partial Shading 3

Fig. 6. I(V) curves of the extended database

We now have a database with a total of 120 maximum power points representative of four health conditions for the PV panel. It should also be noted that these data are representative of the variability inherent in real-world measurements in an outdoor environment. In the next section,

3. Proposed methodology for fault diagnosis

The proposed methodology for faults diagnosis can be described in four main steps, as presented in the flowchart displayed in Fig. 7:

- *The modelling*: For this step, we have chosen a data-driven model based on the 120 maximum power point coordinates extracted from the I (V) curves.
- *Data pre-processing*: This step is decisive for the performance level of the diagnostic method. It consists in choosing the field of information, the descriptive variables, and in implementing any technique allowing to keep the information useful and to eliminate nuisances. In our study, we used data from the maximum power point in the time domain, to which we applied the logarithmic function as in [22].
- *Features extraction and analysis*: The third and fourth steps aim to extract from the maximum power points, the features (signatures) to be analyzed to detect, isolate, and estimate the fault severity. One of the difficulties lies in choosing the most efficient tools for the extraction and analysis of signatures, especially since in our case the data varies with the environmental conditions. This sensitivity makes it difficult to distinguish the healthy state from the faulty ones. Therefore, to ensure reliable fault detection, the most appropriate tool should have the highest ability to separate the data from different health status. In this work, we have selected PCA and LDA for features' extraction and analysis to perform the PS detection and classification.

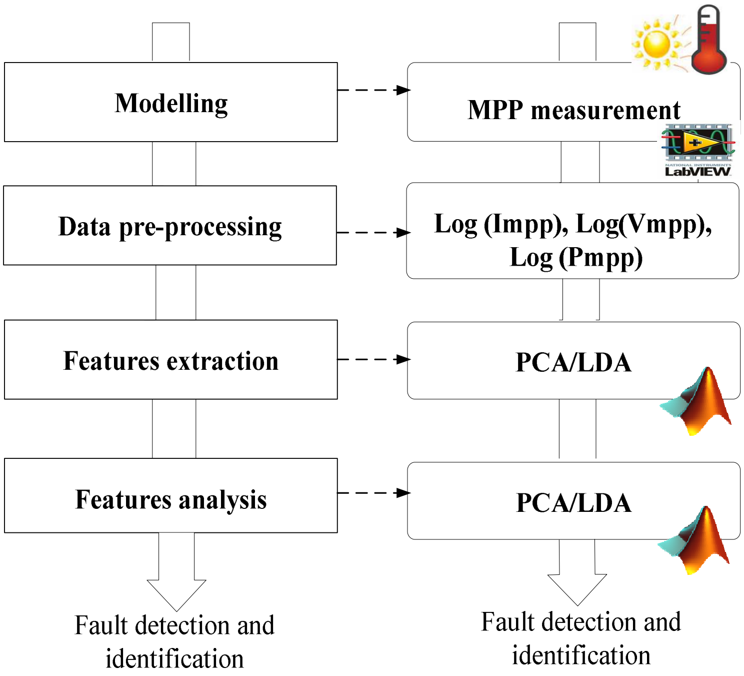


Fig. 7. Flowchart of the four-step diagnosis methodology

3.1.Review of normalization methods in the field of PV

When dealing with maximum power point (MPP) in outdoor conditions, it might not be easy to distinguish healthy status from shading ones. The maximum power point's coordinates change rapidly due to variations in environmental conditions, and it becomes then difficult to differentiate whether a fault or climatic changes cause the MPP variation. In other words, the faulty PV system may have similar MPP’s coordinates as a healthy one. Besides, the fluctuation of MPPs for different fault types may be the same. This is confirmed by the evolutions of the maximum power points’ coordinates plotted in Fig. 8. For the current at MPP, I_{mpp}, we can observe the overlapping of several points in healthy conditions, partial shading 1, and partial shading 2. For the power at the MPP, P_{mpp}, the overlapping is observed in healthy conditions, partial shading 1, and partial shading 3. These overlaps can be expected to be more pronounced when the data acquisition interval is greater. Therefore, to ensure reliable fault detection, one should find appropriate processing techniques that highlight the data's separability.

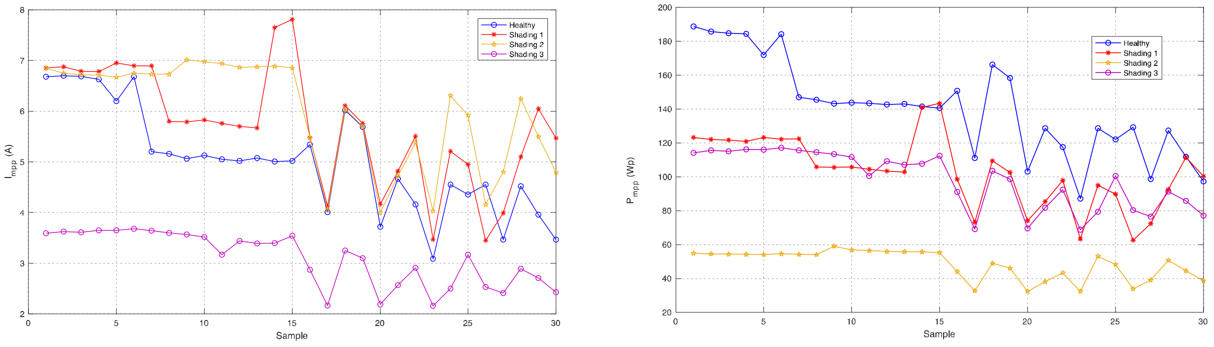


Fig. 8. Overlapping MPP’s coordinates over a wide range of temperature and irradiance

Several methods of normalization have been proposed in the literature to mitigate the effect of the environmental changes. Authors in [3] have defined two normalized parameters. The first one is the ratio of MPP’s voltage by the reference open circuit voltage multiplied by N_{mod}, the number of series-connected PV modules. The second is the ratio of MPP’s current by the reference short-circuit current I_{sc_ref} multiplied by N_{Str} the number of parallel strings in the array. This technique is easy to integrate within a PV inverter whatever its topology. However, it requires the use of two additional modules identical to the PV modules under test. They are considered as modules of reference to measure instantaneous I_{sc_ref} and V_{oc_ref} with respect to solar irradiance and module temperature. This normalization is efficient for detecting and classifying line-to-line and open-circuit fault, but it cannot be used for partial shading. However, the main issue with this method is to guarantee that the modules under test, and the modules of reference are in the same conditions of temperature and irradiance. Another normalization method is proposed in [33]. The PV array current, and voltage, are divided by the STC short-circuit current, and the STC open-circuit voltage, respectively to calculate the normalized parameters, used to determine the entropy of the PV array output power. The method is fast and effective for fault detection, however fault classification has not been considered. A per-unit standardization has been applied in two recent studies [29], [40]. This normalization requires accurate measurements of temperature and irradiance. It also assumes that the temperature coefficients of voltage and current are constant. Nevertheless, several experiments on the PV modules installed at the French observatory SIRTA have revealed that these coefficients vary depending on the meteorological data. Based on this normalization, the diagnosis method in [40] achieved high fault identification percentages for all the studied faults except for partial shading. In that case, the rate of correct identification is 53.66%, which is quite low.

Table 2 summarizes the normalization methods discussed above and the data size used to evaluate each diagnosis approach.

Table 2. Normalized expressions used for PV system diagnosis to mitigate the climatic conditions effect

Variable	Normalized expression	Size of PV data	Reference
Normalized current and voltage at MPP	$I_{mppn} = \frac{I_{mpp}}{N_{str} \times I_{sc\_ref}}$ $V_{mppn} = \frac{V_{mpp}}{N_{\mathrm{mod}} \times V_{oc\_réf}}$	Simulation test : 5576 MPP data 1.43% used for training Experimental test: 5000 MPP data 1% used for training	[3]
Normalized current and voltage at MPP	$I_{mppn} = \frac{I_{mpp}}{I_{sc0}}$ $V_{mppn} = \frac{V_{mpp}}{V_{oc0}}$	No training step	[33]

Per unit normalization of Current and voltage at MPP, ideality factor, open-circuit voltage and short-circuit current.	$I_{mppn}=\frac{I_{mpp}}{I_{mpp0}}\frac{G}{1000}[1+\alpha\frac{I_{mpp0}}{I_{sc0}}(T-25)]$ $V_{mppn}=\frac{V_{mpp}}{V_{mpp0}}[1+\beta\frac{V_{mpp0}}{V_{oc0}}(T-25)]$ $I_{scn}=\frac{I_{sc}}{I_{sc0}}\frac{G}{1000}[1+\alpha(T-25)]$ $V_{ocn}=\frac{V_{oc}}{V_{oc0}}[1+\beta(T-25)]$	123 I-V data 123/3 used for training	[40]
Per unit normalization of Current and voltage at MPP, ideality factor, open-circuit voltage and short-circuit current.	$n_n=\frac{n}{n_0}[1+\alpha(T-25)]$	Simulation test : 4800 I-V data 75% used for training Experimental test : 2500 I-V data 75% used for training	[29]

3.2.Feature Extraction and Analysis with Principle Component Analysis

In our case, before applying PCA, the three variables, I_{mpp}, V_{mpp}, and P_{mmp} are centred (zero mean), and reduced (unit variance). The challenge is also to deal with a limited number of samples (thirty points per health status). The robustness of our proposal to T_{PV} and G_{POA} variations will be evaluated.

3.2.1. Principle

The details of PCA are given in [22]. Briefly, PCA is a multivariate unsupervised technique that uses a linear combination of ***m*** original variables to create a low-dimensional set of new variables named principal components (PCs). The new variables are uncorrelated, and the new dataset is considered optimal when the variance of the information is maximal despite the reduction in the size of the representation space. The original data is arranged in a matrix $X_{[N\times m]}$ where ***N*** represents the number of measurements/ observations for each of the ***m*** variables. The observations include data on healthy and faulty operating conditions. In the following, the PCA algorithm is applied to the matrix $X = \log[V_{mpp} \quad I_{mpp} \quad P_{mpp}]$.

3.2.2. Results in the training stage with the PCA algorithm

In the training step, 80 observations representing 67% of the dataset is used. The PCA is applied to the matrix $X_{[80\times 3]} = \log[V_{mpp} \quad I_{mpp} \quad P_{mpp}]$. We obtain three principal components PCi (i = 1,2,3), among which PC1 and PC2 preserve 99.9% of the initial total variance. The data can be projected in four different sub-spaces spanned by PC1-PC2, PC1-PC3, PC2-PC3, or PC1-PC2-PC3 respectively. The projections of the training data are displayed in Fig.9. Except in the subspace spanned by PC2 and PC3 (there is less information in PC2, and PC3 compared to PC1), the four classes can be clearly distinguished in the other three subspaces: one healthy class C0 and three faulty ones noted C1, C2, and C3 (Table 3). The dispersion of the operating points found in the same group is due to the climatic conditions variations. Table 4 summarizes the percentages of classification, and shows that all the data samples are correctly classified. In this table, we also present the classification performance when using all the points of the I(V) curve in [22]. Despite the variations of T_{PV} and G_{POA}, and using only MPP's coordinates, the classification obtained in the present study is more accurate with 100% rate of classification.

Table 3. The four health status classes

Test condition	Class
Healthy	C_0
Shading 1	C_1
Shading 2	C_2
Shading 3	C_3

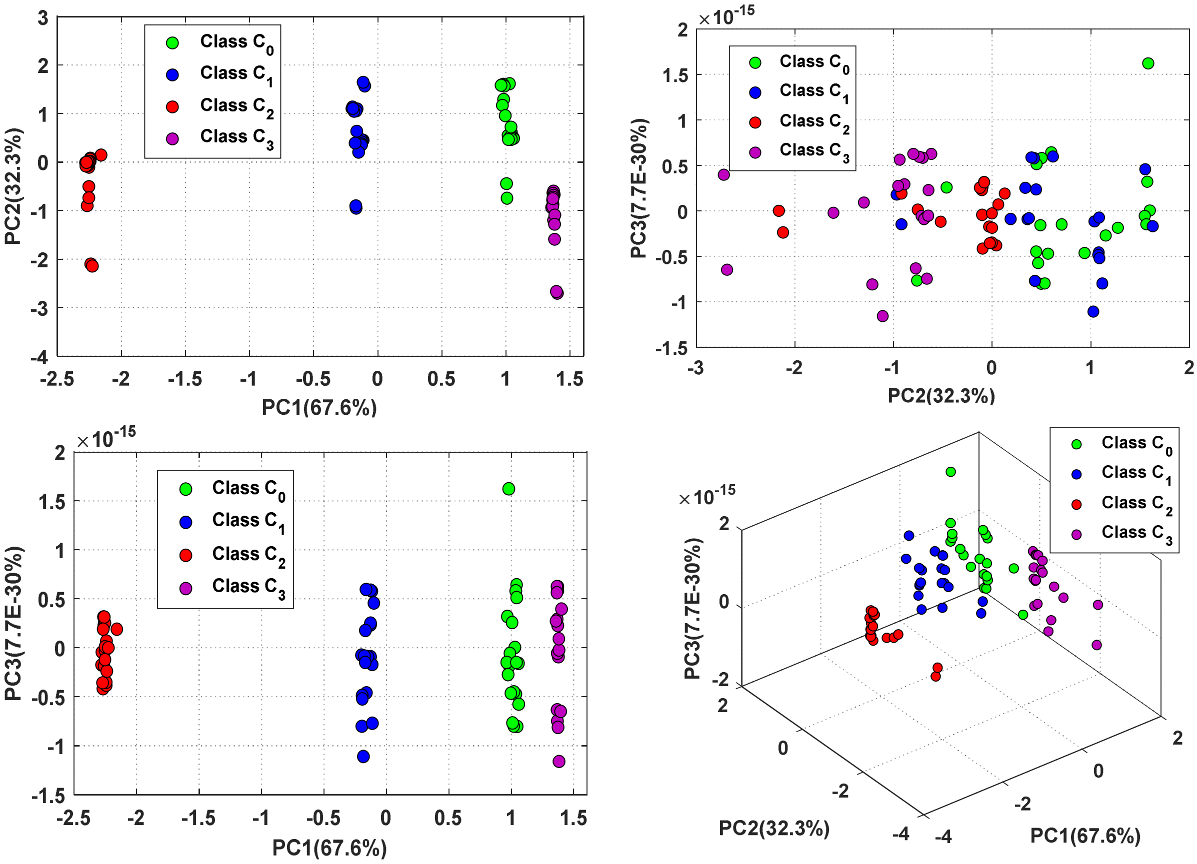


Fig. 9. Projection of the training Maximum Power Points' coordinates in the PCA subspaces

Table 4. Confusion matrix for training dataset classification

A priori Class	A posteriori class							
	Class C_0		Class C_1		Class C_2		Class C_3	
	MPP	I (V)	MPP	I (V)	MPP	I (V)	MPP	I (V)
Class C_0	**100**	97.03	0	0	0	2.97	0	0
Class C_1	0	0	**100**	98.52	0	1.48	0	0
Class C_2	0	12.62	0	0	**100**	87.38	0	0
Class C_3	0	0	0	0	0	0	**100**	100

3.2.3. Results in the validation stage with the PCA algorithm

The remaining 33% of the database is used to test the effectiveness of the trained PCA model. The test matrix is $X_{[40\times3]} = \log[V_{mpp} \quad I_{mpp} \quad P_{mpp}]$ where 10 test samples are used for each health status. The projections of the data in the subspaces spanned by the principal components obtained in the training stage are displayed in Fig.10. The classes related to partial shading 1, 2, and 3 are clearly separated from the healthy one with a classification rate of 100%. These results are also compared with those obtained when using all the points of the curve I(V) (Table 5). They confirm that the analysis of the maximum power point's coordinates with the PCA is effective to detect and classify the partial shading (PS) accurately.

The justification for this high classification rate lies in the fact that partial shading more strongly affects the point of maximum power, which is at the inflexion point of the I(V) curve. Therefore, using only this information makes the classification more effective than using all the points of the I(V) curve.

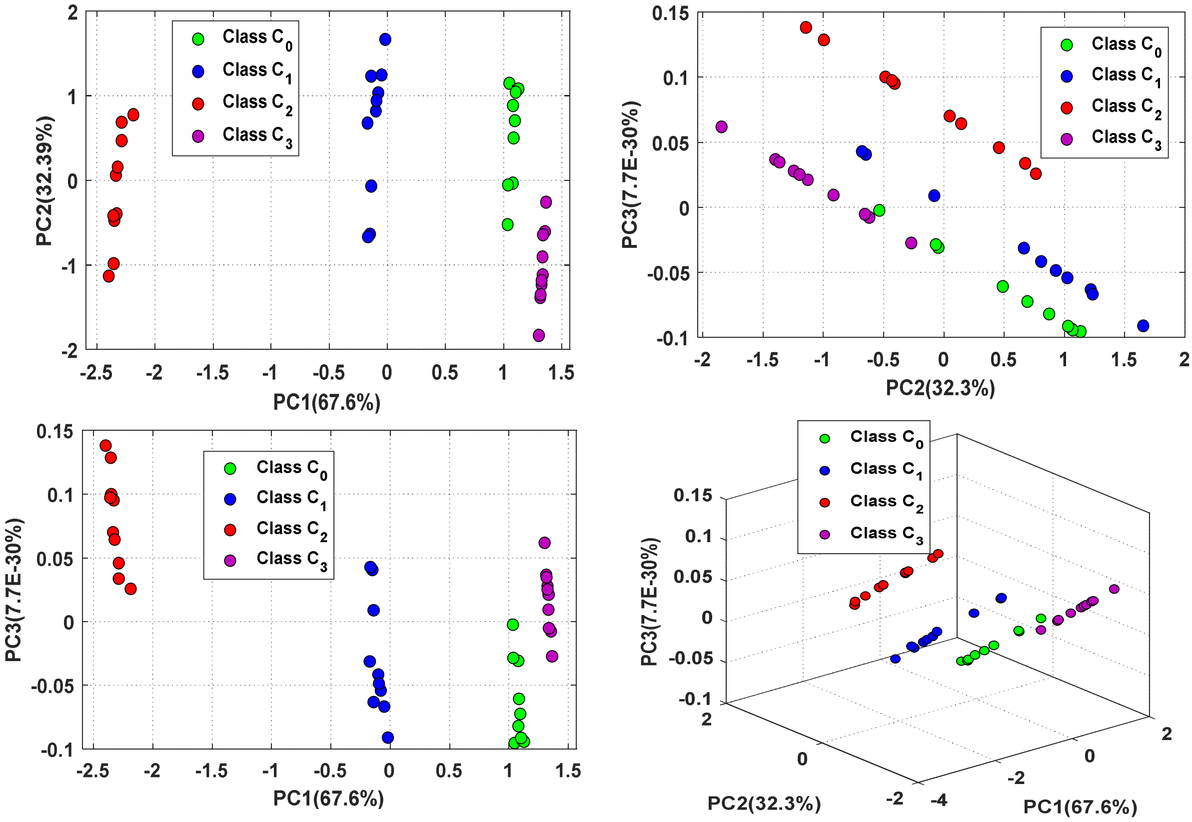


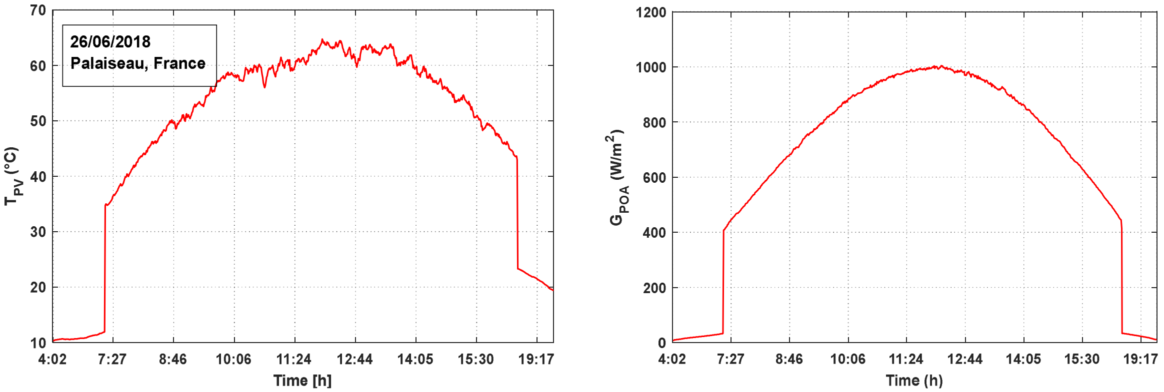
Fig. 10. Projection of the test data in the PCA subspaces

Table 5. Confusion matrix for test data classification

A priori class	A posteriori class							
	Class C_0		Class C_1		Class C_2		Class C_3	
	MPP	I(V)	MPP	I(V)	MPP	I(V)	MPP	I(V)
Class C_0	**100**	97.03	0	0	0	2.97	0	0
Class C_1	0	0	**100**	99.01	0	0.99	0	0
Class C_2	0	0	0	0	**100**	100	0	0
Class C_3	0	0	0	0	0	0	**100**	100

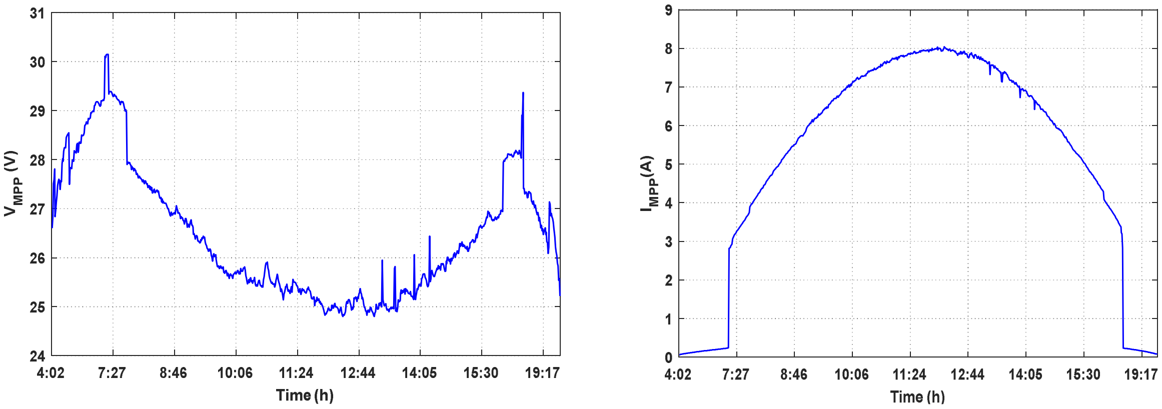
3.2.4. *Robustness of the classification*

This section aims to validate the robustness of the fault diagnosis method based on PCA and maximum power point (MPP) coordinates. To do so, previously used faulty MPP data (collected on March 21, 2019), are merged with MPP data collected during healthy operation on June 26, 2018. The corresponding daily profiles of temperature and solar irradiance, as well as the MPP voltage and current evolution under these conditions, are plotted in Fig.11. The robustness is evaluated using the data extracted at different times, and for moderate and high irradiation values. Table 6 presents the new samples used and the ranges of variation of T_{PV} and G_{POA}.



(a) Temperature of the module

(b) Irradiance



(c) Voltage at MPP

(d) Current at MPP

Fig. 11. Clear day profiles

Table 6. Extracted samples for robustness evaluation

Samples	Time (h)	T_{PV} (°C)	G_{POA} (W/m^2)
[100 ; 109]	[7h26 ; 7h33]	[36.46 ; 37.9]	[445.53 ; 468.46]
[200 ; 209]	[8h44 ; 8h52]	[48.55 ; 49.81]	[680.71 ; 703.64]
[450 ; 459]	[12h03 ; 12h10]	[63.18 ; 64.19]	[992.10 ; 1004.50]
[500 ; 509]	[12h43 ; 12h50]	[59.68 ; 62.49]	[967.88 ; 982.80]
[621 ; 630]	[14h22 ; 14h29]	[56.93 ; 58.4]	[797.53 ; 815.39]
[700 ; 709]	[15h29 ; 15h37]	[48.96 ; 50.93]	[606.28 ; 628.28]

The new matrix with the merged data $X_{[40\times3]} = \log[V_{mpp} \quad I_{mpp} \quad P_{mpp}]$ is projected in the PCA subspace spanned by the two principal components (PC1, PC2) obtained in the learning stage with the data collected on March 21, 2019. The results are presented in Fig.12. We observe that class C_0 formed with the new healthy measurements is clearly separated from the other three classes of data corresponding to partial shading (PS) conditions. A successful classification of 100% is obtained. Moreover, we notice that all the new test data is positioned between classes C_2 and C_3, as observed in the previous validation stage. These results are encouraging and prove the robustness of the method for PS detection and classification of PV modules.

When looking at the maximum power point's coordinates projected in the PCA subspace, one can observe that classes can be linearly separated. Therefore in the following, linear discriminant analysis (LDA) is used to retrieve the classification laws for fault identification.

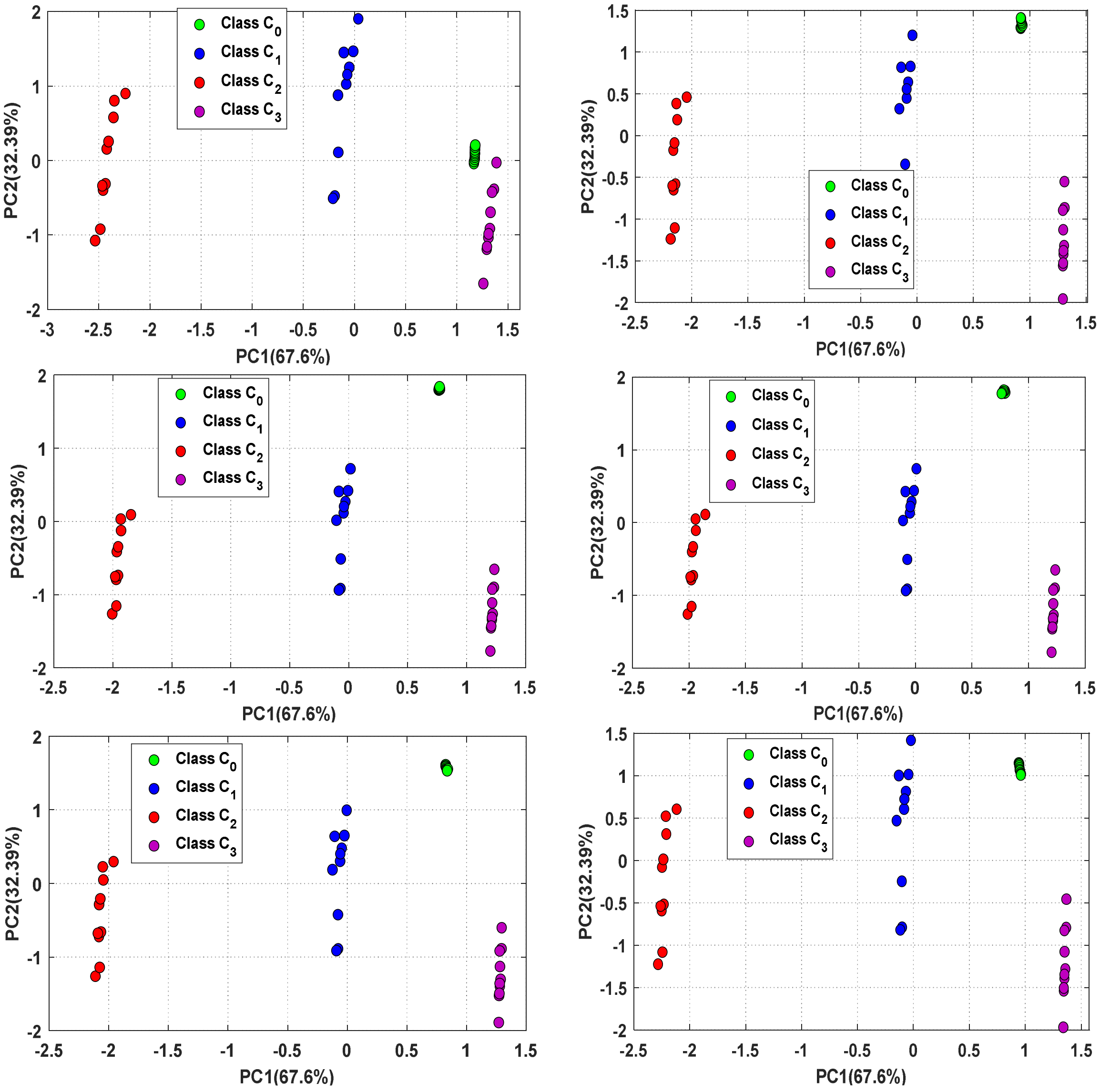


Fig. 12. Projection of the merged dataset in PCA subspace (PC1, PC2)

3.3.Features analysis and extraction with Linear Discriminant Analysis (LDA)

3.3.1. Formulation of the LDA algorithm

The LDA algorithm used in [41] [42] [43] will be extended for application to photovoltaic system diagnosis. LDA and PCA share a common characteristic, which is dimension reduction. Nevertheless, theoretically, LDA is more powerful for data classification than PCA because LDA performs dimensionality reduction while keeping maximum the discriminatory information among the classes. For this reason, LDA is generally used to improve data classification when different data classes overlap with the PCA algorithm.

Unlike the PCA, the LDA is a supervised analysis technique. It involves a learning phase during which c a priori groups are defined with M descriptive variables and N observations or samples distributed into a priori groups, as shown in Figure 13. LDA is both descriptive and decisive. In its descriptive phase, it reduces the size of the data and separates the data according to their a priori defined groups. Linear combinations of the variables in each group create c -1 discriminant axes. Each axis is assigned a percentage of the data variance, which is calculated from the corresponding eigenvalue. As in PCA, the discriminant axes are ordered in descending order of eigenvalues. In its decisive phase, the LDA generates c -1 hyperplanes or separation laws that characterize the differences or similarities between the groups allocated a priori. Thanks to these laws, it is possible to classify new samples of data described by the same explanatory variables without a priori knowledge of their group.

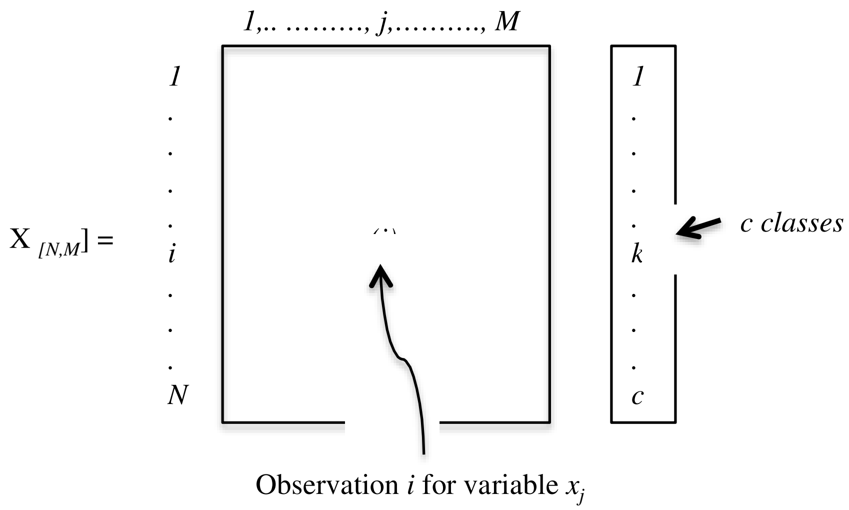


Fig. 13. Database organization for LDA algorithm

To perform best linear combinations, LDA aims to obtain a maximum interclass variance which highlights the differences between the classes, and a minimum intraclass variance [44]. The interclass and intraclass variances are respectively estimated by the between matrix variance S_B and the within matrix variance S_w.

Mathematically formulating, this analysis technique is based on maximizing the following objective function, also called Fisher criterion [45]:

$$J(u) = \frac{u^T S_B u}{u^T S_w u} \tag{1}$$

where u is the linear expression that best discriminates all classes of data.

Given ***N*** observations $(x_1, x_2, ..., x_N)$ of ***M*** centered reduced explicative variables, ***N****ᵢ*, the total number of observations for each class C_i, $i = 1, 2, ..., c$.

Let's consider μ_i as the mean value of observations belonging to C_i and μ as the mean value of the entire dataset given by:

$$\mu_i = \frac{1}{N_i} \sum_{x_n \in C_i} x_n \tag{2}$$

$$\mu = \frac{1}{N} \sum_{n=1}^{N} x_n \tag{3}$$

The between variance matrix can be expressed as:

$$S_B = \sum_{i=1}^{c} (\mu_i - \mu)^T (\mu_i - \mu) \tag{4}$$

The within variance matrix is calculated as:

$$S_w = \sum_{i=1}^{c} S_i \tag{5}$$

where :

$$S_i = \sum_{x_n \in C_i} (x_n - \mu_i)^T (x_n - \mu_i) \tag{6}$$

The eigenvector decomposition of $S_w^{-1} S_B$ gives the discriminant axes and their corresponding eigenvalues $\lambda_1, \lambda_2, ..., \lambda_{c-1}$ that indicate the discriminant power of each axis. The eigenvectors associated with the largest eigenvalues define the discriminant axes that span the LDA space for classification. The original data centered reduced are finally projected onto this space to give the data class separability.

In the decisive phase, LDA uses three quantities to classify the new data samples in one of the defined a priori groups: posterior probability, prior probability, and cost.

The posterior probability that a new instance or observation ***l*,** described by the explicative variables, is classified into the class C_i is estimated by the Bayes formula:

$$\hat{P}(C_i / l) = \frac{\mathrm{P}(l / C_i).\mathrm{P}(C_i)}{\mathrm{P}(l)} \tag{7}$$

where $\mathrm{P}(C_i)$ represents the prior probability of class C_i.

With this probabilistic approach, the unknown observation $\boldsymbol{l}$ to be identified will be assigned to the class C_i with the highest posterior probability and minimum expected classification cost:

$$\hat{y} = \underset{y=1,..,c}{\arg\min}(P(C_1 / k)C(y / C_1) + .. + P(C_i / k)C(y / C_i) + .. + P(C_c / k)C(y / C_c)) \tag{8}$$

Where $\hat{y}$ is the predicted class, c is the number of classes and $C(y / C_i)$ is the cost of classifying an observation as y when its true class is C_i.

3.3.2. Results of the training stage with LDA algorithm

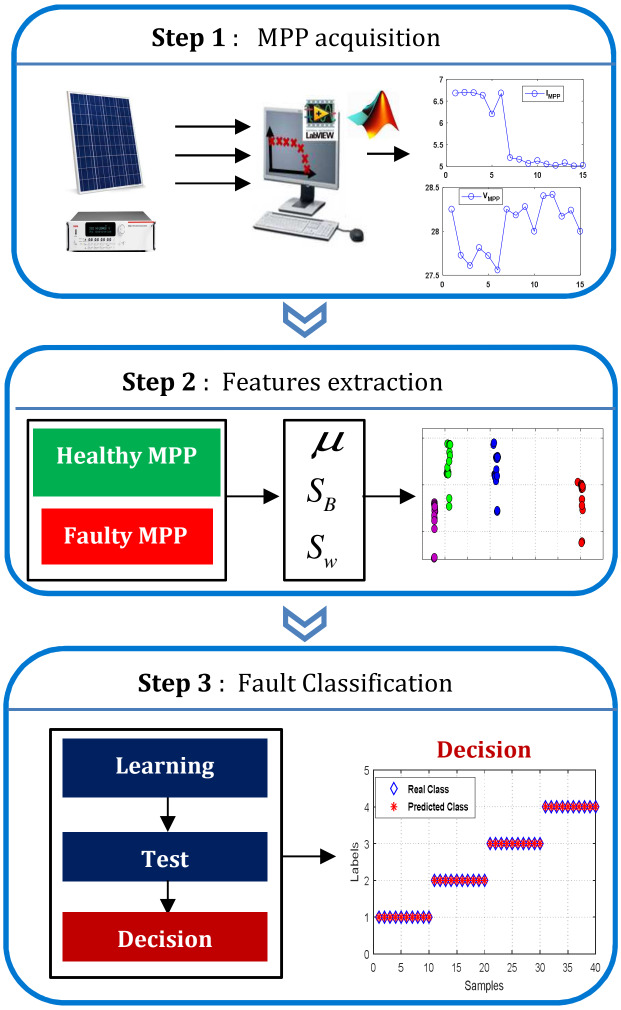


Fig. 14. Fault identification flowchart with LDA algorithm

The flowchart of the proposed algorithm is shown in Fig.14. The LDA is applied to extract the faults discriminating characteristics from maximum power points (MPPs) coordinates. Then, the labels of the data are assigned according to their class of membership (Table 7), and the classifier computes the linear separation laws allowing the classification of new observations. The four a priori classes for original labels are retrieved from the classification results with the PCA algorithm.

Table 7: Fault types and original labels

Condition	Class	Label
Healthy	C_0	1

Partial Shading 1	C_1	2
Partial Shading 2	C_2	3
Partial Shading 3	C_3	4

This training stage is carried out using 67% of the data as original labels. The same matrix $X_{[80\times3]} = \log[V_{mpp} \quad I_{mpp} \quad P_{mpp}]$ built with data recorded on March 21, 2019 is used to extract the fault features. Three discriminant axes, also called Fisher axes, are obtained. Fig.15 shows the projection of the training dataset into the 2 – dimensional LDA subspace spanned by the first and the second discriminant axes (DA1, DA2). These two axes represent 99.7%, and 0.18% of the data variance, respectively. In this figure, we have also drawn the linear boundaries between the classes. The magenta line is the boundary that separates the data for classes C_0 and C_3. On the other hand, the blue line separates the classes C_0 and C_1, and the red line separates the observations of C_1 from those of C_2.

Let us consider ($\boldsymbol{x_1, x_2}$) as the coordinates of an instance $\boldsymbol{l}$ in the LDA subspace (DA1, DA2). The three boundaries are described by three affine functions f_1, f_2 and f_3 where each function f_i is given by:

$$f_i(x_1,x_2) = L_i(1)x_1 + L_i(2)x_2 + K_i, i = 1,2,3 \tag{9}$$

The coefficients L_i and K_i, i=1, 2, 3, are given in Table 8.

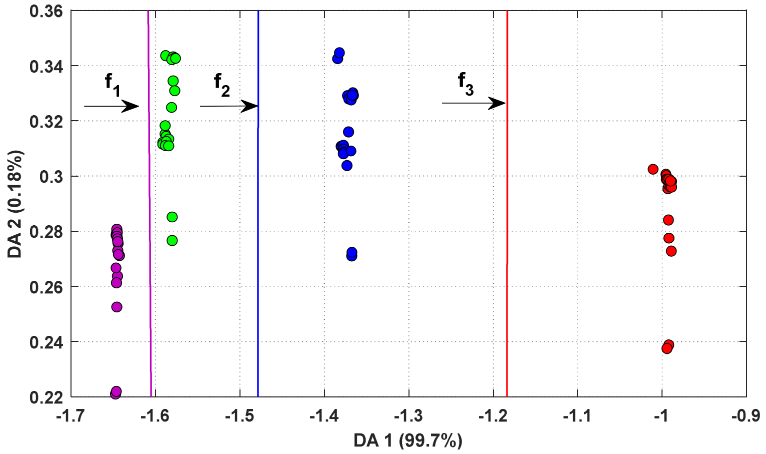


Fig. 15. Projection of the training data in the subspace (DA1, DA2)

Table 8: Linear boundaries coefficients

f_i	$L_i(1)$	$L_i(2)$	K_i
f_1	2.93 10^3	71	4.7 10^3
f_2	-1.82 10^4	59	-2.15 10^4
f_3	-1.01 10^4	17	-1.49 10^4

With the cross-validation method, the same success rate (100%) as for PCA is obtained. LDA also entirely separates the different classes even under climatic variations. In the following, the obtained separation laws will be used to classify new samples; it is the validation stage.

3.3.3. Results of the validation stage with LDA

The validation is done in two steps. In the first one, the calculated separation laws are used to classify the remaining MPP (test data) representing 33% of the database collected on March 21, 2019. The results, displayed in figure 16 show the successful classification of the test data.

In the second step, the separation laws are evaluated with the healthy samples extracted from measurements on June 26, 2018, mixed with 33% of partial shading data measured on March 21, 2019. The classification result is displayed in Fig.17. With the separation laws, all the test samples are correctly classified in their a priori groups; no misclassification errors were found. With the cross-validation method, discrimination's effectiveness between the measurements under variable experimental conditions is proved with a success rate of 100 % for all the classes (Table 9).

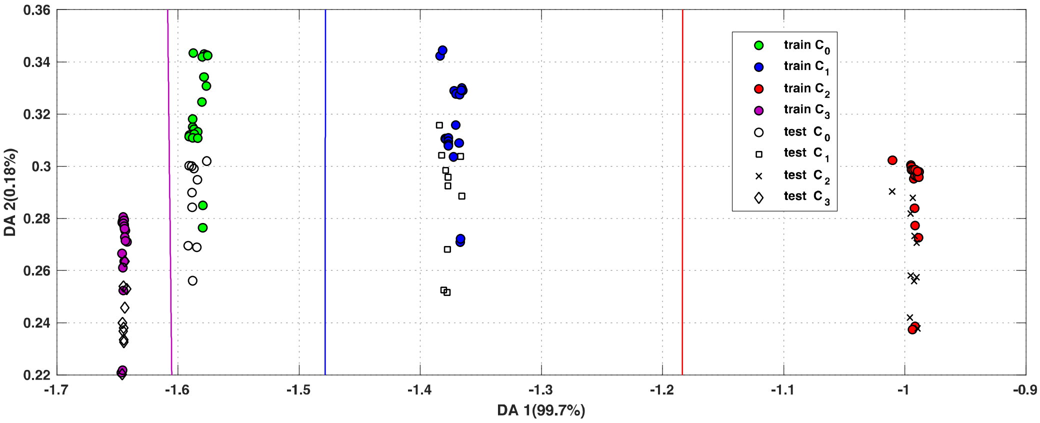
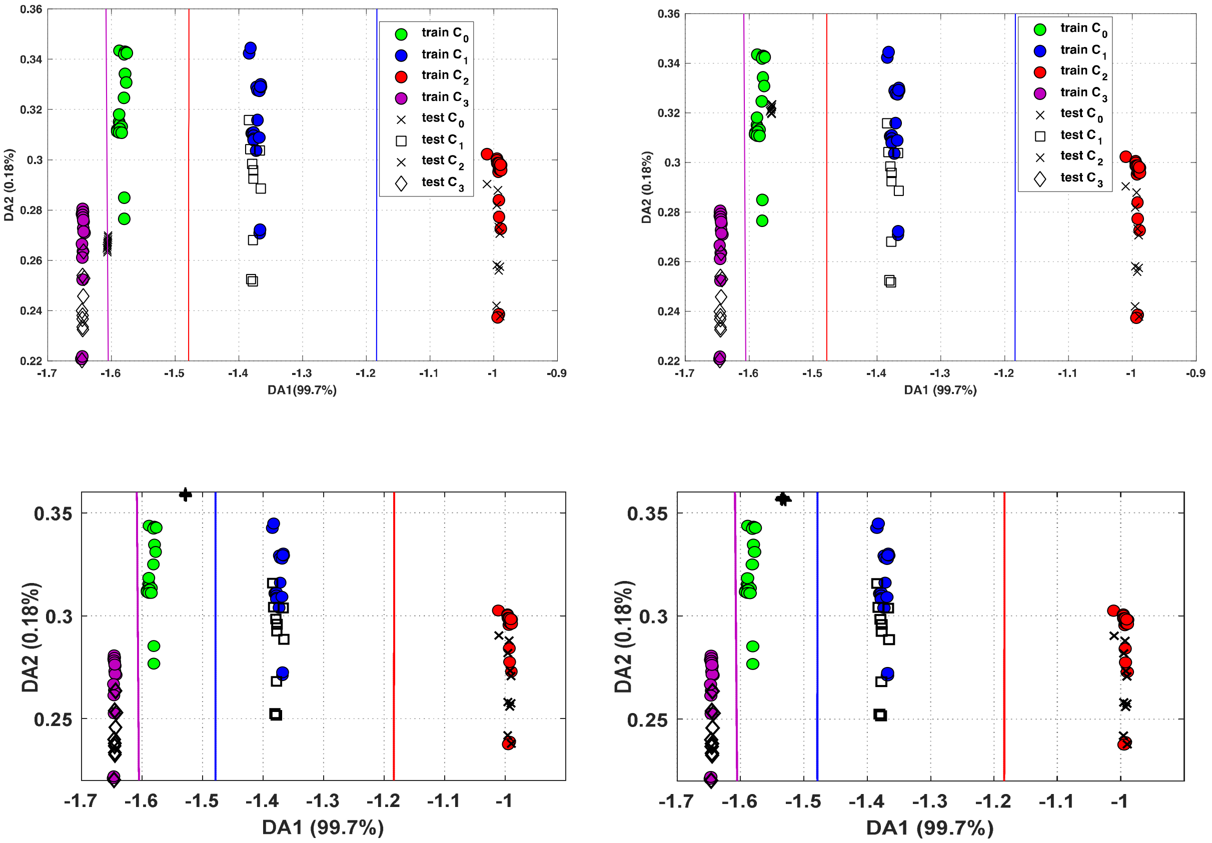


Fig. 16. Classification of test data in the discrimination space (DA1, DA2)



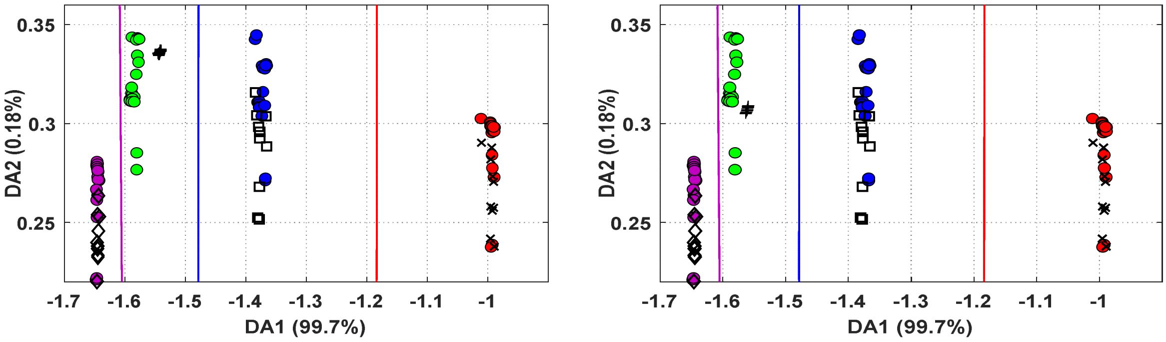


Fig. 17. Classification of mixed data in the discrimination space (DA1, DA2)

Table 9. Confusion matrix for test dataset with LDA

Class a priori	Class a posteriori			
	Class C_0 (%)	Class C_1 (%)	Class C_2 (%)	Class C_3 (%)
Class C_0	**100**	0	0	0
Class C_1	0	**100**	0	0
Class C_2	0	0	**100**	0
Class C_3	0	0	0	**100**

Even if for the studied cases the obtained identification is correct, it can be observed in Fig.17 that healthy state's data corresponding to samples 'test C_0', lie on the separation boundary between classes C_0 and C_3. However, as the probability of assignment to class C_0 is higher, these samples are predicted with label 1 and correctly identified as healthy data. Nevertheless, this ambiguity underlines the fact that in order to have good performance in classifying several types of defects with different levels of severity, it will be beneficial to have a rich database. It must contain several configurations (operating points and states of health) under different environmental conditions.

We can conclude that at the PV module level, the classification of measurements through LDA is successful as long as a large amount of labelled data is available for offline learning.

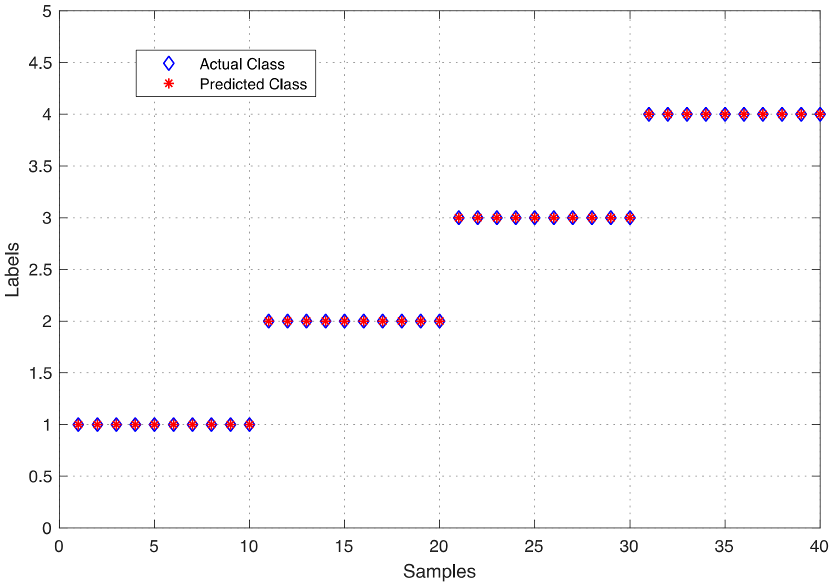


Fig. 18. Healthy and faulty class prediction based on LDA

4. Discussion and comparison

The classification results obtained through the analysis of the maximum power point's (MPP) coordinates with PCA and LDA for partial shading diagnosis are very close. Table 10 summarizes the main advantages and their limitations of the diagnosis methods proposed in this paper, the methods developed in our previous work, especially for large PV plants.

Fault diagnosis based on the analysis of maximum power point's (MPP) coordinates appears to be relevant for large PV plants. Indeed, this approach enables diagnosing a large number of PV modules in real-time faster than using all the points on the I(V) curve, thus reducing power losses and costs.

From a field-application perspective, it would be of interest to evaluate these methods with other faults.

MPP-based methods have up to now only been evaluated at the PV module level. It would be interesting to evaluate these methods at the PV string and array levels.

Besides the ability to PV fault diagnosis, the proposed methods can also be used to quantify the power loss in real-time, thus allowing maintenance optimization. Moreover, they can be combined with reconfiguration algorithms to limit the shading effect. When several PV modules are connected in series, the shading fault detection and notably the isolation is decisive to balance the modules or at least mitigate the power production gaps between strings. Different reconfigurable architectures have been developed, taking into account the shading condition and have allowed significant energy savings [46], [47].

Table 10: Overview of the proposed methods for partial shading detection in PV systems

Pre-processing	**Advantages**	**Limitations for large scale PV plants**
▪ $\log(\frac{\boldsymbol{v}}{\eta})$ ▪ $\log(\boldsymbol{i})$ ▪ $\log(\frac{\boldsymbol{p}}{\eta})$	- Adapted with online I (V) tracer - Sensitive to fault severity - Suitable for small-scale PV systems	- Less convenient - Time-consuming and functional disturbance - Power losses - Additionnal cost with local I (V) tracers - Sensitive to I (V) curve quality
▪ $\log(I_{MPP})$ ▪ $\log(V_{MPP})$ ▪ $\log(P_{MPP})$	- Earlier fault detection allowing an optimized maintenance - Cost-effective : no additional measurement of temperature or irradiance for data normalization - Saving memory storage - Sensitive to fault severity - Robust to climatic variations - Possible power losses quantification (**LDA**)	- Efficient and accurate fault identification depend on the learning data set (size and variety) (**LDA**)

Method	Variables
- **I (V) curve analysis – PCA** [22]	- Current ***i*** - Voltage ***v*** - Power ***p*** **303** observations for each variable
- **MPP analysis - PCA**	- Current $\boldsymbol{I_{MPP}}$ - Voltage $\boldsymbol{V_{MPP}}$ - Power $\boldsymbol{P_{\mathrm{MPP}}}$ **30** observations for each variable
- **MPP analysis - LDA**	

5. Conclusion

Fault detection and identification in PV systems are essential to expedite the maintenance and minimize the power losses. This study has presented a cost-effective and straightforward data-driven diagnosis approach to detect and identify partial shading using principal component analysis (PCA) and linear discriminant analysis (LDA). No prior information on the PV system and no other sensing circuit are needed. The proposed methods take advantage of available measurements in PV systems: voltage, current, and power at the maximum power point (MPP). The effectiveness of the proposed methods has been successfully proved at the PV module level. Based on collected data under real working conditions (variable irradiation and temperature), we have demonstrated that the methods can correctly detect and classify the partial shading fault with a success rate of 100%. The methods are also easy to integrate into a monitoring system for real-time diagnosis and are especially convenient for large-scale PV plants. The data required for the learning stage can benefit from the vast amount of data available in PV plants under operation for decades. Otherwise, the learning database could be generated from simulation models. It would be of interest to evaluate the proposed methodology with other faults (open-circuit, short-circuit, and contact degradation).

Acknowledgement

The authors gratefully recognize the financial support of the Ministry of Higher Education and Scientific Research in Tunisia and Paris-Saclay University who provided the scholarship to the PhD student.

References

[1] A. Golnas, "PV System Reliability: An Operator's Perspective," *IEEE J. Photovolt*, vol. 3, no. 1, pp. 416–421, 2013, doi: 10.1109/JPHOTOV.2012.2215015.
[2] IEA, "Reliability Study of Grid Connected PV Systems," T7-08, 2002.
[3] Y. Zhao, R. Ball, J. Mosesian, J. de Palma, and B. Lehman, "Graph-Based Semi-supervised Learning for Fault Detection and Classification in Solar Photovoltaic Arrays," *IEEE Trans. Power Electron*, vol. 30, no. 5, pp. 2848–2858, 2015, doi: 10.1109/TPEL.2014.2364203.
[4] D. C. Jordan, T. J. Silverman, J. H. Wohlgmuth, S. R. Kurtz, and K. T. VanSant, "Photovoltaic failure and degradation modes," *Prog. Photovolt: Res. Appl.*, pp. 318–326, 2017, doi: 10.1002/pip.2866.
[5] IEA, "Assessment of Photovoltaic Module Failures in the Field," 2017.
[6] X. Sun, R. V. K. Chavali, and M. A. Alam, "Real-time monitoring and diagnosis of photovoltaic system degradation only using maximum power point—the Suns-Vmp method," *Prog Photovolt Res Appl*, vol. 27, no. 1, pp. 55–66, 2019, doi: 10.1002/pip.3043.
[7] M. Bressan, Y. El Basri, A. G. Galeano, and C. Alonso, "A shadow fault detection method based on the standard error analysis of I-V curves," *Renew.Energy*, vol. 99, pp. 1181–1190, 2016, doi: 10.1016/j.renene.2016.08.028.
[8] A. Y. Appiah, X. Zhang, K. A. Ben Beklisi, and F. Kyeremeh, "Review and performance evaluation of photovoltaic array fault detection and diagnosis techniques," *Int. J. Photoenergy*, vol. 2019, pp. 1–19, 2019, doi: https://doi.org/10.1155/2019/6953530.
[9] A. Triki-Lahiani, A. Bennani-Ben Abdelghani, and I. Slama-Belkhodja, "Fault detection and monitoring systems for photovoltaic installations: A review," *Renew. Sustain. Energy Rev*, vol. 82, pp. 2680–2692, 2018, doi: https://doi.org/10.1016/j.rser.2017.09.101.
[10] D. S.Pillai, F. Blaabjerg, and N. Rajasekar, "A comparative Evaluation of Advanced Fault Detection Approaches for PV Systems," *IEEE Journal of Photovoltaics*, vol. 9, no. 2, pp. 513–527, 2019, doi: 10.1109/JPHOTOV.2019.2892189.
[11] S. Daliento *et al.*, "Monitoring, Diagnosis, and Power Forecasting for Photovoltaic Fields: A Review," *International Journal of Photoenergy*, pp. 1–13, 2017, doi: https://doi.org/10.1155/2017/1356851.
[12] F. Spertino *et al.*, "A power and energy procedure in operating photovoltaic systems to quantify the losses according to the causes," *Sol.Energy*, vol. 118, pp. 313–326, 2015, doi: https://doi.org/10.1016/j.solener.2015.05.033.
[13] IEA, "Review on Infrared and Electroluminescence Imaging for PV Field Applications," T13-10, 2018.

[14] M. Wolfgang, C. Hirschl, L. Neumaier, A. Kolger, and A. Plosch, “Identifikation von fehlerhaften PV Modulen und Anlangen mit Thermografie,” presented at the 29. Symposium Photovoltaische Solarenergie 2014, Deutschland, 2014, pp. 1–11, doi: 10.13140/2.1.3636.2881.

[15] J. A. Tsanakas, L. D. Ha, and F. AL Shakarchi, “Advanced inspection of photovoltaic installations by aerial triangulation and terrestrial georeferencing of thermal/visual imagery,” *Renew.Energy*, vol. 102, pp. 224–233, 2017, doi: https://doi.org/10.1016/j.renene.2016.10.046.

[16] S. Roy, M. K. Alam, F. Khan, J. Johnson, and J. Flicker, “An Irradiance-Independent, Robust Ground-FaultDetection Scheme for PV Arrays Based on SpreadSpectrum Time-Domain Reflectometry (SSTDR),” *IEEE Trans. Power Electron*, vol. 33, no. 8, pp. 7046–7057, 2018, doi: 10.1109/TPEL.2017.2755592.

[17] M. K. Alam, F. Khan, J. Johnson, and J. Flicker, “PV ground-fault detection using spread spectrum time domain reflectometry (SSTDR),” presented at the IEEE Energy Conversion Congress and Exposition (ECCE), USA, 2013, pp. 1015–1020, doi: 10.1109/ECCE.2014.6953848.

[18] M. K. Alam, F. Khan, J. Johnson, and J. Flicker, “PV arc-fault detection using spread spectrum time domain reflectometry (SSTDR),” presented at the IEEE Energy Conversion Congress and Exposition (ECCE), USA, 2014, pp. 3294–3300, doi: 10.1109/ECCE.2014.6953848.

[19] T. Takashima, J. Yamaguchi, and M. Ishida, “Fault detection by signal response in PV module strings,” presented at the 33rd IEEE Photovoltaic Specialists Conference, USA, 2008, doi: Fault detection by signal response in PV module strings.

[20] T. Takashima, J. Yamaguchi, K. Otani, T. Oozeki, K. Kato, and M. Ishida, “Experimental studies of fault location in PV module strings,” *Sol. Energy Mater. and Sol. Cells*, vol. 93, no. 6–7, pp. 1079–1082, 2009, doi: https://doi.org/10.1016/j.solmat.2008.11.060.

[21] S. Fadhel, A. Migan, C. Delpha, D. Diallo, I. Bahri, M. Trabelsi, M.F. Mimouni, “Data-driven Approach for Isolated PV Shading Fault Diagnosis Based on Experimental I-V curves Analysis,” presented at the IEEE International Conference on Industrial Technology (ICIT), France, 2018, pp. 927–932, doi: 10.1109/ICIT.2018.8352302.

[22] S. Fadhel, C. Delpha, D. Diallo, I. Bahri, A. Migan, M. Trabelsi, M.F. Mimouni, “PV shading fault detection and classification based on I-V curve using principal component analysis: Application to isolated PV system,” *Sol.Energy*, vol. 179, pp. 1–10, 2019.

[23] Y. El Basri, M. Bressan, L. Seguir, H. Alawadhi, and C. Alonso, “A proposed graphical electrical signatures supervision method to study PV module failures,” *Sol.Energy*, vol. 116, pp. 247–256, 2015, doi: https://doi.org/10.1016/j.solener.2015.02.048.

[24] A. K. Tossa, Y. M. Soro, Y. Azoumah, and D. Yamegueu, “A new approach to estimate the performance and energy productivity of photovoltaic modules in real operating conditions,” *Sol.Energy*, vol. 110, pp. 543–560, 2014, doi: 10.1016/j.solener.2014.09.043.

[25] M. Merchaoui, A. Sakly, and M. F. Mimouni, “Particle swarm optimisation with adaptive mutation strategy for photovoltaic solar cell/module parameter extraction,” *Energy Conv. and Manag*, vol. 175, pp. 151–163, 2018, doi: https://doi.org/10.1016/j.enconman.2018.08.081.

[26] Y. Chaibi, M. Malvoni, A. Chouder, M. Boussetta, and M. Salhi, “Simple and efficient approach to detect and diagnose electrical faults and partial shading in photovoltaic systems,” *Energy Conversion and Management*, vol. 196, pp. 330–343, 2019, doi: https://doi.org/10.1016/j.enconman.2019.05.086.

[27] D. S.Pillai and N. Rajasekar, “Metaheuristic algorithms for PV parameter identification: A comprehensive review with an application to threshold setting for fault detection in PV systems,” *Renew. Sustain. Energy Rev*, vol. 82, pp. 3503–3525, 2018, doi: https://doi.org/10.1016/j.rser.2017.10.107 Get.

[28] S. Kichou, S. Silvestre, L. Guglielminotti, L. Mora-Lopez, and E. Munoz-Ceron, “Comparison of two PV array models for the simulation of PV systems using five different algorithms for the

parameters identification," *Renew.Energy*, vol. 99, pp. 270–279, 2016, doi: https://doi.org/10.1016/j.renene.2016.07.002.

[29] Z. Chen, L. Wu, S. Cheng, P. Lin, W. Yue, and W. Lin, "Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I-V characteristics," *Applied Energy*, vol. 204, pp. 912–931, 2017, doi: https://doi.org/10.1016/j.apenergy.2017.05.034.

[30] S. Spataru, D. Sera, T. Kerekes, and R. Teodorescu, "Diagnostic method for photovoltaic systems based on light I–V measurements," *Sol.Energy*, vol. 119, pp. 29–44, 2015, doi: 10.1016/j.solener.2015.06.020.

[31] J. D. Bastidas-Rodr´ıguez, E. Franco, G. Petrone, C. A. Ramos-Paja, and G. Spagnuolo, "Model Based Degradation Analysis Of Photovoltaic Modules Through Series Resistance Estimation," *IEEE Trans. Indus. Electr.*, vol. 62, no. 11, pp. 7256–7265, 2015, doi: 10.1109/TIE.2015.2459380.

[32] D. Wang, "A method for instantaneous measurement of PV I-V characteristics and its application for MPPT control," presented at the 35th IEEE Photovoltaic Specialists Conference, USA, 2010, pp. 2904–2907, doi: 10.1109/PVSC.2010.5614505.

[33] A. Khoshnami and I. Sadeghkhani, "Sample entropy-based fault detection for photovoltaic arrays," *IET Renew. Power Gener*, vol. 12, no. 16, pp. 1966–1976, 2018, doi: 10.1049/iet-rpg.2018.5220.

[34] M. Dhimish, V. Holmes, B. Mehrdadi, and M. Dales, "Multi-layer photovoltaic fault detection algorithm," *IET High Voltage*, vol. 4, pp. 244–252, 2017, doi: 10.1049/hve.2017.0044.

[35] F. Harrou, A. Dairi, B. Taghezouit, and Y. Sun, "An unsupervised monitoring procedure for detecting anomalies in photovoltaic systems using a one-class Support Vector Machine," *Sol.Energy*, vol. 179, pp. 48–58, 2019, doi: https://doi.org/10.1016/j.solener.2018.12.045.

[36] E. Garoudja, F. Harrou, Y. Sun, K. Kara, A. Chouder, and S. Silvestre, "Statistical fault detection in photovoltaic systems," *Sol.Energy*, vol. 150, pp. 485–499, 2017, doi: https://doi.org/10.1016/j.solener.2017.04.043.

[37] L. Garcia-Gutierrez, M. Bressan, F. Jimenez, S. de-las Heras, and C. Alonso, "Design of a Global Maximum Power Point Tracking (GMPPT) for PV array based on precise PV shadow model," presented at the 7th International Conference on Renewable Energy Resarch and Applications (ICRERA), Paris, France, Oct. 2018, pp. 275–280, doi: 10.1109/ICRERA.2018.8566880.

[38] M. Aquib, S. Jain, and V. Agarwal, "A Time-Based Global Maximum Power Point Tracking Technique for PV System," *IEEE Trans. Power Electr.*, vol. 35, no. 1, pp. 393–402, 2020, doi: 10.1109/TPEL.2019.2915774.

[39] K. Kobayashi, I. Takano, and Y. Sawada, "A study of a two stage maximum power point tracking control of a photovoltaic system under partially shaded insolation conditions," *Sol. Energy Mater. and Sol. Cells*, vol. 90, pp. 2975–2988, 2006, doi: https://doi.org/10.1016/j.solmat.2006.06.050.

[40] H. Zhu, L. Lu, J. Yao, S. Dai, and Y. Hu, "Fault diagnosis approach for photovoltaic arrays based on unsupervised sample clustering and probabilistic neural network model," *Sol.Energy*, vol. 176, pp. 395–405, 2018, doi: https://doi.org/10.1016/j.solener.2018.10.054.

[41] M. Benbouzid, C. Delpha, Z. Khatir, S. Lefebvre, and D. Diallo, "Faults Detection and Diagnosis in a Static Converter," in *Electrical Machine Diagnosis*, ISTE., 2011, pp. 272–319.

[42] D. Y. Jung, S. M. Lee, H. Wang, J. H. Kim, and S. H. Lee, "Fault detection method with PCA and LDA and its application to induction motor," *J. Cent. South Univ. Technol.*, vol. 17, pp. 1238–1242, 2010, doi: https://doi.org/10.1007/s11771-010-0625-y.

[43] R. J. Martis, U. R. Acharya, and L. C. Min, "ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform," *Biomedical Signal Processing and Control*, vol. 8, pp. 437–448, 2013, doi: https://doi.org/10.1016/j.bspc.2013.01.005.

[44] A. Quinquis, *Digital Signal Processing using MATLAB*, ISTE. André Quinquis, 2008.

[45] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*. New York, 1979.

[46] M. A. Chaaban *et al.*, “Adaptive photovoltaic system,” presented at the 36th Annual Conference on IEEE Industrial Electronics Society (IECON), USA, 2010, pp. 3192–3197, doi: 10.1109/IECON.2010.5675047.

[47] M. Alahmad, M. A. Chaaban, S. kit Lau, J. Shi, and J. Neal, “An adaptive utility interactive photovoltaic system based on a flexible switch matrix to optimize performance in real-time,” *Sol.Energy*, vol. 86, pp. 951–963, 2012, doi: https://doi.org/10.1016/j.solener.2011.12.028.