Engineering Science and Technology, an International Journal 24 (2021) 1168–1179

HOSTED BY

Contents lists available at ScienceDirect

Engineering Science and Technology, an International Journal

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

Discriminative common vector in sufficient data Case: A fault detection and classification application on photovoltaic arrays

Yasemin Onal *, Umit Cigdem Turhal

Bilecik Seyh Edebali University, Department of Electrical and Electronics Engineering, Bilecik 11000, Turkey

ARTICLE INFO

Article history:
Received 22 September 2020
Revised 13 February 2021
Accepted 25 February 2021
Available online 14 March 2021

Keywords:
Photovoltaic array fault
Fault detection and classification
Sufficient data case
Discriminative common vector
Signal processing

ABSTRACT

In this study, the derivation of the Discriminative Common Vector (DCV) approach which is first introduced for a face recognition task in the insufficient data case, for the sufficient data case is obtained and it is applied for a photovoltaic (PV) panel fault detection and classification. Two experimental studies are performed including two different fault configurations. In the first experimental study, as the faulty conditions open-circuit, short-circuit, and partial shading conditions are taken and healthy condition is taken as reference. Thus, a four-class fault detection and classification scheme is constructed. In the second experimental study, the serial resistance degradation fault is considered. This fault detection and classification scheme includes four classes that are healthy and three different serial resistance degradation. The data used in the experimental studies are formed to be 1x3 dimensional vectors which include the current, voltage, and power values obtained from the simulations in the PSIM program. In all two experimental studies for each class, a discriminative common vector (DCV) which represents the common properties of that class, thus, having a high discriminative ability is obtained. As a contribution to the literature, the derivation of DCVA which has high discrimination ability for sufficient data case, and usage of it for PV panels fault detection and classification is proposed for the first time in this study. The proposed method's performance is evaluated with the performance of PCA method that is recently used for the fault detection and classification problem in PV panel systems in the literature. In the first experimental study, the proposed method's performance (99%) is obtained significantly higher than the performance of the PCA method (95%). And in the second experimental study, while PCA can only detect the faulty condition but cannot classify the serial resistance degradation, the proposed method can both detect and classify with 99% accuracy the PV panel serial resistance degradation.

1. Introduction

The number of studies on renewable energy systems has increased in recent years due to factors such as insufficient fossil resources, environmental damage, and increases in energy consumption. Among renewable energy systems, photovoltaic (PV) energy systems are more preferred because being a source of natural energy, they are sustainable, and they can be easily installed anywhere [1].

Since PV energy systems operate in a natural environment, they are affected by environmental factors. This causes complex and various faults. These faults include aging, shadow, hot spot, short circuit, open circuit fault, and partial shading fault. Faults in PV panels affect the reliability, efficiency, service life, and stable operation of the energy system negatively and even cause fire hazards. The study shows that the annual energy losses due to faults were 3.6% for the first year of operation, 6.6% for the second year, and 18.9% for the first year of operation in site B [2]. Due to the loss of energy, the monitoring and fault diagnosis of PV panels are becoming more and more critical [3–4]. The fault diagnosis methods of the PV arrays include signal processing methods [5–7], I-V curves model-based diagnosis methods [8–12], and intelligent diagnosis methods [13–29].

The signal processing methods consist of the earth capacitance method, time-domain reflectometry TDR method, multi-sensor method, analytical method, and satellite observations method. In the earth capacitance method, open circuit fault is determined by using the earth capacitance value in series-connected PV arrays [5]. TDR method is used to detect wiring anomalies in arrays, including interruptions in the circuit, insulation defects, open circuit, and inversed polarity. This method is basically based on

* Corresponding author.
E-mail addresses: yasemin.onal@bilecik.edu.tr (Y. Onal), ucigdem.turhal@bilecik.edu.tr (U.C. Turhal).
Peer review under responsibility of Karabuk University.

https://doi.org/10.1016/j.jestch.2021.02.017
2215-0986/

waveform analysis of the output voltage [6]. However, PV faults are detected by turning off the PV arrays in the earth capacity and TDR methods. Faults cannot be detected in real-time using these methods. In the multi-sensor method, data from multiple sensors placed in the PV array is analyzed. In the study, the number of open and short circuits of PV arrays is tried to be determined by analyzing the irradiance level, temperature, number of wires of PV arrays, and output power [7]. PV faults can be detected in real-time using the multi-sensor method. However, this method is expensive since it requires additional equipment to be placed on PV arrays.

The I-V curve model-based diagnostic methods use the PV array model to determine the type and location of the fault occurring in the PV arrays [8–9]. In model-based diagnostic methods, faults of PV arrays are accurately detected without the need for any additional equipments [10]. However, it is problematic to create and implement an accurate model of the large-scale PV array system in this method. In the study, R_s and R_{sh} are calculated from the I-V curve properties and array degradation is determined by analyzing such indicators. An 11% degradation is noted over more than 20 years [11]. In the study, Kalman filters are used to determine the output power lowering in PV arrays by the I-V relationship [12].

Intelligent diagnosis methods or statistical pattern recognition methods identify faulty operating conditions in PV arrays using particle swarm optimization PSO algorithm, support vector machine SVM algorithm, an artificial neural network ANN algorithm, fuzzy and neuro-fuzzy classification algorithm, and statistical pattern recognition. In refference [13], the PSO algorithm is proposed to determine all possible local maximum voltage points according to the P-U characteristic to locate the starting position of the agents. In this study, a SVM based fault diagnosis algorithm is presented to detect short circuit, open circuit, and partial shading faults occurring in PV arrays. Short circuit current, open circuit voltage, maximum power current, and maximum power voltage are selected as input parameters of the SVM algorithm [16–17]. The ANN algorithm is used for diagnosing faults in a PV, using power device failure, overheating of power devices, very low output voltage value, and abnormal voltage meter values in references [20–22]. In reference [23], a fuzzy classification algorithm based on the analysis of theoretical curves of a PV system is used. Under a certain operating condition, using a number of properties such as voltage ratio and power ratio, the temperature data of radiation and PV arrays are simulated using LabVIEW software. In reference [24], a fault detection and classification algorithm using a neuro-fuzzy classifier is used in case of a serial fault, defect bypass diode, and blocking diode faults. Maximum power and open circuit voltage deviations are selected as input parameters. However, due to the uncertain number of nodes in the hidden layer in ANN, partial learning, over-fitting learning, and under-fitting learning become difficult to implement. Therefore, ANN-based intelligent diagnosis algorithms cannot accurately predict the faults that occur in PV arrays. In recent years, fault diagnosis use of PCA method is popular in PV system [25–28]. In reference [28], PCA is used for fault diagnosis in case of a partial shading condition in a PV system. In reference [29], PCA is used for dimension reduction purposes, again in a PV fault diagnosis system as a preprocessing step before SVM.

In this study, derivation of discriminative common vector (DCVA) approach in case of sufficient data is proposed based on reference [30] and a common vector in the first stage is obtained that is unique for each class although the null space of the within-class scatter matrix of all classes approaches to zero. As a very important renewable energy source, the PV array system is an important example corresponding to sufficient data case, thus, this derivation is used in a fault detection and classification task for PV array. The discriminative common vector approach (DCVA) [31] is a statistical pattern recognition method based on the common vector approach (CVA) [32], which is originally presented for a face recognition problem in case of insufficient data. In both methods, the main idea is to find a common vector for each class. But there is a difference between them: DCVA uses the within-class scatter matrix of samples of all classes while CVA uses the within-class scatter matrix of only the samples in that class. Both methods are defined for the insufficient data case. However, Gulmezoglu et al. obtained the derivation of the CVA method for the sufficient data case, and they theoretically defined the common vector that is unique for each class in case of sufficient data [30].

In this paper, as the PV array fault detection and classification method, the novel proposed derivation of the DCVA method for sufficient data case is used. As far as we know, according to the literature, this is the first study. In the experimental studies, as the DCVA is a statistical pattern recognition algorithm, its performance is evaluated with PCA, also a statistical pattern recognition algorithm. The data to be used in the application consists of the voltage, current and power values obtained as a result of the simulation of the array models produced using PSIM. The results obtained in the experimental studies are given in comparison with the results of the PCA method, which has come to the fore in the literature, especially in recent years. According to the results obtained, the proposed method provided significantly better classification accuracies reaching 99% for the two of the experimental studies while it is obtained 95% accuracy in PCA for the first experimental study. However, in the second experimental study, PCA is failed. This is because PCA is a dimension reduction method and it eliminates the useful information required to discriminate the different classes from each other [33].

The contributions of this paper can be summarized as follows: First, derivation of DCVA in case of sufficient data which is novel in the literature is proposed. Second, DCVA that keeps the specific and distinctive information for a sample class is used in a PV system fault diagnosis for the first time in the literature. The proposed study is applied offline, for the case of a PV system. Third, the proposed method for PV fault diagnosis only uses voltage, current and power values of a PV system, so, its sensitivity to environmental factors is low. In addition, it performs the detection and the classification independently of any PV sizes. Fourth, the cost of determining and classifying PV array fault does not increase with the algorithm proposed. Fifth, the DCVA model is simpler than the models in the literature and results in good and clear data apparency. The proposed method's performance is evaluated with the performance of the PCA method that is recently used for the fault detection and classification problem in the PV array system in the literature. The proposed method separates healthy data from faulty data. It makes a good classification in open circuit, short circuit, and partial shading compared to other classification methods in the literature. Especially in series resistance degradation fault, while PCA cannot separate the faulty conditions, CVA provides an extremely high separability.

This paper is organized as follows: The proposed fault detection and classification method in case of sufficient data are explained in section II. Photovoltaic system and type of PV faults in PSIM are given in Section III. The experimental results for PV array fault detection and classification using DCVA and PCA are given in Section IV. Finally, brief conclusions are drawn in Section V.

2. Proposed fault detection and classification method

2.1. CVA and DCVA

The CVA first emerged as a solution to the problem of isolated word recognition when the sample dimension for each class was equal to or greater than the number of samples in the training set [32]. This method is based on the principle of finding a common

vector which is single and constant for each class. For any class, class common vector is obtained by projecting any sample from that class onto the indifference subspace, where the common properties of the data belonging to the same class are defined. This vector is then used in recognition. The common vector for each class is calculated, using the eigenvectors that span the null space obtained via the Eigen analysis of the within-class scatter matrix of only the samples belonging to that class. Let $x_1, x_2, \cdots., x_m$ be the word vectors that exist in the training set for each word class and are supposed to be linear independent. In this case, each word vector, x_j is defined as

$$x_j = x_{j,dif} + x_{com} + \varepsilon_j, \;\; j = 1, 2, \cdots., m \tag{1}$$

where $x_{j,dif}$ represents the differences caused by various factors for any word class, x_{com} represents the common aspects of different samples in the word class and ε_j represents the error vector. The within-class scatter matrix for any word class, say c can be found as in Eq. (2) in CVA method

$$S_W = \sum_{j=1}^{m} (x_j - \mu)(x_j - \mu)^T \tag{2}$$

where $\mu = \frac{1}{m}\sum_{j=1}^{m} x_j$, is the mean vector of the word class c.

In the DCVA, the common vector is calculated by the within-class scatter matrix of the samples of all classes instead of using the within-class scatter matrix of the samples of a single class. The within-class scatter matrix used in DCVA is as given in Eq. (3)

$$S_W = \sum_{i=1}^{C} \sum_{j=1}^{m} \left(x_j^i - \mu_i\right)\left(x_j^i - \mu_i\right)^T \tag{3}$$

where $\mu_i = \frac{1}{m}\sum_{j=1}^{m} x_j^i, i = 1, \cdots, C$. C is the total class number.

The common vector for CVA and DCVA is obtained using the null space obtained from the Eigen analysis of the scatter matrices given in Eqs. (2) and (3), respectively. Both methods are defined for insufficient data case. In other words, in both methods the within-class scatter matrices S_W have a null space. Let it be assumed that $R^d$is the original sample space, and V and $V^{\perp}$, denote the range space and the null space of the within-class scatter matrices, S_W, given in Eqs. (2) and (3). Then, R^dcan be defined as $R^d = V \oplus V^{\perp}$ where

$$V = \text{span}\{\alpha_k | S_W \alpha_k \neq 0, \;\; k = 1, \cdots, r\} \tag{4}$$

$$V^{\perp} = \text{span}\{\alpha_k | S_W \alpha_k = 0, \;\; k = r+1, \cdots, d\} \tag{5}$$

As $r < d$ to be the rank of matrix S_W given in Eqs. (2) and (3), $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$ is the set of the orthonormal eigenvectors. These eigenvectors correspond to the eigenvalues which are non-zero of the S_W matrices and $\{\alpha_{r+1}, \alpha_{r+2}, \cdots, \alpha_d\}$ is the set of the orthonormal eigenvectors that correspond to the eigenvalues which are zero of the S_W matrices. Then, $Q = [\alpha_1 \;\; \cdots \;\; \alpha_r]$ and $\overline{Q} = [\alpha_{r+1} \;\; \cdots \;\; \alpha_d]$ are the matrices used to define the difference and the indifference subspaces, respectively. At this stage, as $P = QQ^T$ and $\overline{P} = \overline{Q}\,\overline{Q}^T$ are the orthogonal projection matrices onto the V and $V^{\perp}$, respectively, components of any sample $x_j^i \in R^d$ as given in Eq. (1) has a single form of decomposition such that

$$x_{j,dif}^i = Px_j^i = QQ^T x_j^i \in V \tag{6}$$

$$x_{com}^i = \overline{P}x_j^i = \overline{Q}\,\overline{Q}^T x_j^i \in V^{\perp} \tag{7}$$

Here, $x_{j,dif}^i$ is the projection of the sample x_j^i onto the difference subspace of the sample space and x_{com}^i is the projection of the sample x_j^i onto the indifference subspace of the sample space, and x_{com}^i can be written as [31]

$$x_{com}^i = x_j^i - x_{j,dif}^i - \varepsilon_j^i = x_j^i - Px_j^i - \varepsilon_j^i \tag{8}$$

The x_{com}^i given in Eq. (8) is called the common vector. In DCVA the obtained common vectors for each class are used to define the scatter of the class common vectors, S_{com}. In recognition, vectors that maximize S_{com} are found and used to obtain a discriminative common vector for each class.

$$S_{com} = \sum_{i=1}^{C} (x_{com}^i - \mu_{com})(x_{com}^i - \mu_{com})^T \tag{9}$$

where $\mu_{com} = \frac{1}{C}\sum_{i=1}^{C} x_{com}^i$ is the mean vector of the class common vectors. The W matrix, consisting of orthonormal optimal projection vectors that maximize the distribution in Eq. (9), provides the following criteria

$$J(W_{opt}) = \arg\max_W \|W^T S_{com} W\| \tag{10}$$

Then, using the W matrix, discriminative common vectors for each class are obtained.

As can be seen, CVA and DCVA are both methods that obtain the common vector using the null space of the sample space. In other words, they can be applied in case of insufficient data($d \geq m$). Otherwise, if ($d < m$), CVA and DCVA cannot be applied in this way. Because, the null space of the sample space approaches to zero. However, Gulmezoglu et al. obtained the derivation of CVA method, in case of sufficient data in their study and they defined a common vector that is single and constant for each class [30]. In the following section a similar derivation is proposed for DCVA, in this paper. Thus, the usage of DCVA in sufficient data case has been possible.

2.2. Derivation of DCV in sufficient data case

A sample in the R^d sample space is defined as given in Eq. (1). In this case, the common vector for a sample class in the CVA method is the only solution found by minimizing the F metric given below with respect to x_{com} [30].

$$F = \frac{1}{2}\sum_{j=1}^{m} \|\varepsilon_j\|^2 = \frac{1}{2}\|x_j - x_{j,diff} - x_{com}\|^2 \tag{11}$$

Here, m is the total number of samples in a sample class and $\|\varepsilon_j\|$ is the Euclidean norm of the vector ε_i. If we adapt this metric to DCVA method, the following metric is obtained:

$$J(W_{opt}) = \arg\max_W \|W^T S_{com} W\| \tag{12}$$

Thus, in the DCVA method, the common vector for any sample class would be the only solution found by minimization of the F metric in Eq. (12) with respect to $x_{com}^i, i = 1, 2, .., C$. As result of this minimization, the common vectors for each class can be obtained as follows:

$$x_{com}^i = x_j^i - x_{j,diff}^i \;\; \forall j = 1, \cdots, m, i = 1, \cdots, C \tag{13}$$

where $x_{j,diff}^i = \langle x_j^i, \alpha_1\rangle\alpha_1 + \cdots .. + \langle x_j^i, \alpha_r\rangle\alpha_r$ and the $\{\alpha_1, \alpha_2, \cdots, \alpha_r\}$ vector set is the set given in Eq. (4).

The explanations given so far for DCVA are those for insufficient data cases where the sample size is greater than the number of samples. The null space of all the samples within-class scatter matrix that appeared in case of insufficient data is the space that is used to obtain the common vector. In the case of sufficient data, the null space approaches to zero, but not completely equals to zero. Thus, all the eigenvectors of the within-class scatter matrix has an eigenvalue that is not zero. The smallest one approaches to zero but not equals to zero. But yet, the indifference subspace can also be defined for sufficient data case. The ability to define

the indifference subspace can be achieved by choosing the k value satisfy the Eq. (14) [34]. Then, the difference and the indifference subspaces of the within-class scatter matrix obtained for all classes can be defined using the orthonormal base vectors. These orthonormal base vectors $u_j \epsilon R^d, j = 1, 2, .., k-1 (k-1 < d)$ and $u_j \epsilon R^d, j = k, k+1, \cdots, d$, span the difference and indifference subspaces, respectively. They are chosen to satisfy the following criteria such that

$$\left(\sum_{j=k}^{d} \lambda_j\right) / \left(\sum_{j=1}^{d} \lambda_j\right) < L \tag{14}$$

For the value of L in Eq. (14) it has been indicated that 5% gives good performance [35]. In this case, $x^i_{j,diff}$ and x^i_{com} given in Eq. (6), Eq. (7) respectively for sufficient data case can be written as follows:

$$x^i_{j,diff} = Px^i_j = \sum_{j=1}^{k-1} \langle x^i_j, u_j \rangle u_j \tag{15}$$

$$x^i_{com} = P^{\perp} x^i_{com} = \sum_{j=k}^{d} \langle x^i_{com}, u_j \rangle u_j \tag{16}$$

Here $P = \sum_{j=1}^{k} QQ^T$, $P^{\perp} = \sum_{j=1}^{k} \overline{QQ}^T$ are the orthogonal projection matrices, where $Q = [u_1 \quad \cdots \quad u_{k-1}]$ and $\overline{Q} = [u_k \quad \cdots \quad u_d]$.

Under these assumptions the F metric given in Eq.12 can be written as:

$$F = \frac{1}{2} \sum_{i=1}^{C} \sum_{j=1}^{m} \|x^i_j - Px^i_j - P^{\perp} x^i_{com}\|^2 = \frac{1}{2} \sum_{i=1}^{C} \sum_{j=1}^{m} \|P^{\perp}(x^i_j - x^i_{com})\|^2 \tag{17}$$

The F metric given in Eq. (17) is a multivariable function in which the variable number is equal to the total class number that is $x^i_{com}, i = 1, \cdots, C$. In order to obtain the common matrix for each class, the F metric is minimized each time with respect to one of the variables (x^c_{com}). Then, x^i_{com} is obtained as follows:

$$x^i_{com} = P^{\perp} x^i_{ave} = P^{\perp} \left(\frac{1}{m} \sum_{j=1}^{m} x^i_j\right), i = 1, .., C \tag{18}$$

If Eq. (18) is written in Eq. (17) F metric becomes

$$F = \frac{1}{2} \sum_{j=k}^{d} u_j^T S_W u_j \tag{19}$$

From the minimization of F according to u_j under the constraint $\|u_j\| = 1$for $j = k, .., d$, u_j base vectors become eigenvectors of S_W indifference subspaces. In this case, S_W

$$S_W = \sum_{j=1}^{m} (x^i_j - x^i_{ave})(x^i_j - x^i_{ave})^T \tag{20}$$

After minimization ofF

$$F_{min} = \frac{1}{2} \sum_{j=k}^{d} U_j^T S_W U_j = \frac{1}{2} (\lambda_k + \lambda_{k+1} + \cdots .. + \lambda_d) \tag{21}$$

Here, $(\lambda_k, \lambda_{k+1}, \cdots .., \lambda_d)$ are the smallest eigenvalues of S_W and $(u_k, u_{k+1}, \cdots .., u_d)$ are the eigenvectors corresponding to these eigenvalues, respectively. The common vector for each class can be obtained with the projection of the average vector of that class onto the indifference subspace $V^{\perp}$.

3. Photovoltaic system modeling and type of PV faults

3.1. Modeling of PV system

PV array manufacturer's catalog includes open-circuit voltage, short-circuit current, voltage at maximum power point (MPP), current at MPP, and maximum power values at standard test condition (STC). The standard test condition is defined as the condition where the irradiation value equals to 1000 W/m^2 at temperature 25 °C. Along with these parameters, the temperature coefficients in V_{oc} and I_{sc} can be obtained from the catalog information.

Due to the existence of different array types in the PV array industry, it is important to consider different parameters in the equivalent circuit to be used in real operating conditions, for accurate and reliable modeling and simulation of the PV array systems. To simulate the operation of a PV array, the researchers used a single diode model [36]. Initially, the values of the resistance used are obtained from the catalog information of the manufacturers. In addition, ideality factor, saturation current, and photon current are needed for these values [37].

In this study, a PV array is modeled using a single diode model with five parameters. The current equation of the PV array can be obtained as shown in Eq. (22) [38–39].

$$I = I_{ph} - I_{so}\left(e^{\frac{V + R_s.I}{V_t n_s}} - 1\right) - \frac{V + R_s.I}{R_{sh}} \tag{22}$$

The output current is defined by using the parameters, such that are the photon current produced in STC (I_{ph}), the saturation current (I_{so}), the serial and the parallel array resistances (R_s ,R_{sh}), the number of connected cells to the array (n_s) and the thermal voltage (V_t). The thermal voltage is calculated using the Boltzmann's constant K (1.38×10^{-23} J/K), the ideal diode factor A, the cell temperature of PV T, the electron charge q ($1.602176565 \times 10^{-19}$). Then, the thermal voltage is expressed as in Eq. (23).

$$V_t = \frac{AKT}{q} \tag{23}$$

The photon current I_{ph} is defined by using the short-circuit current I_{sc}, the current temperature coefficient K_i, light intensity S, and T and it is expressed as in Eq. (24).

$$I_{ph} = (I_{sc} + K_i(T - 298.15))S/100 \tag{24}$$

The saturation current I_{so} is defined by the current temperature coefficient, the cell temperature, the short-circuit current, the electron charge, the voltage temperature coefficient K_v, the open-circuit voltage V_{oc}, ideal diode factor, and Boltzmann's constant as in Eq. (25).

$$I_{so} = (I_{sc} + K_i(T - 298.15)) / \left(e^{q\left(\frac{V_{oc} + K_v(T - 298.15)}{AKTn_s}\right)} - 1\right) \tag{25}$$

The PV array parameters selected in the modeling are chosen to increase the reliability and quality of modeling fault conditions. For this reason, current, voltage and power parameters are obtained from the output of the PV array. The PV array used in this study is a STS-156P-255 W module and the array parameters are shown in Table 1. In the following, we simulated a PV array with the same specifications as the one considered in the literature [16].

In the physical model of the PV array, PSIM which can be used for solar irradiation and ambient temperature changes is used. The physical model used is shown in Fig. 1. Many parameter inputs are required for the physical model. Some of these parameters can be obtained from the catalog information, while other parameters such as photo current, diode saturation current, series and shunt resistors, and ideality factor must be obtained by means of trial and error through the physical model.

Table 1
Parameters of PV array at STC.

Parameters (STC)	Model	Value
Maximum power, P_{max}	STS-156P-255 W	255 W
Current at P_{max}, I_{mpp}		8.32A
Voltage at P_{max}, V_{mpp}		30.65 V
Short-circuit current, I_{sc}		8.95 A
Open-circuit voltage, V_{oc}		37.8 V

3.2. Modeling of PV faults

PV arrays constitute the most important part of PV power systems. However, PV arrays can be easily affected by environmental factors, so different types of faults occur in PV arrays. These faults are partial shading, aging and heating, open-circuit, short-circuit, crack and the series resistance degradation. When the weather is dark or cloudy, shades occur on the PV arrays, resulting in partial shading fault because the PV arrays cannot receive the illumination on the shaded area. Aging occurs in the PV arrays after a certain period of time. Aged arrays cannot fully absorb the sunlight and this leads to the decrease of the amount of the irradiation that PV array can receive. The aging fault in PV arrays is indicated as partial shading fault. The crack fault of PV array is indicated as partial shading fault because it will have a serious scattering effect on the sunlight.

Figs. 2 and 3 give the comparison of I-V curves and P-V curves obtained from the PV array in case of standard test condition, shading fault condition, and partial shading fault condition. In case of standard test condition, current and power curves are obtained by applying the temperature value 25 °C and the irradiation value 1000W/m^2. In case of partial shading fault, current and power curves are obtained by applying the irradiation value 400W/m^2. In case of partial shading fault, first 20 sub-cells were shaded

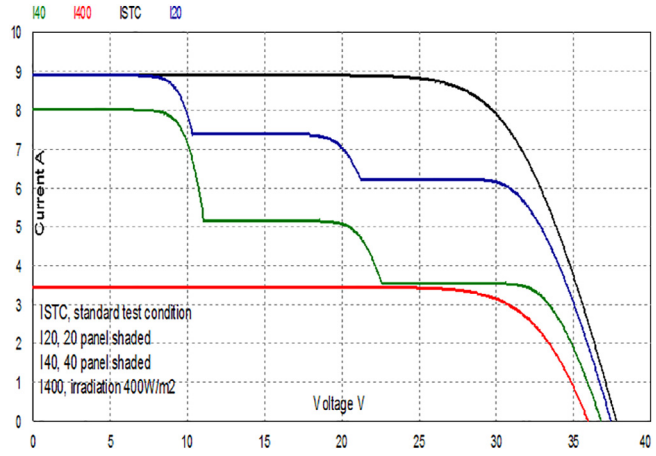

Fig. 2. Comparison of I-V curves obtained from different faults in PV array.

and then, 40 more sub-cells were shaded in the study. Three sub-cell are equipped with three bypass diodes; each one is connected in anti-parallel to protect a PV sub-cell. Multiple peaks occur in all I-V curves depending on the type of shading applied in the shading fault condition. These peaks indicate that the efficiency of the PV array is reduced since the maximum power is reduced under shading fault. In Fig. 2, it is seen that the I-V curves obtained when the shaded PV cells are formed in the area between the curve formed in the STC condition and the curve formed when the partial shading fault to 400 W/m^2. The I-V curve shows that the partial shading fault and the irradiance change fault have the same characteristic of I–V curve. Partial shading, aging, cracking of PV arrays is actually a decrease in sunlight, this fault can be shown as partial shading faults. As a result, faults occurring in PV arrays

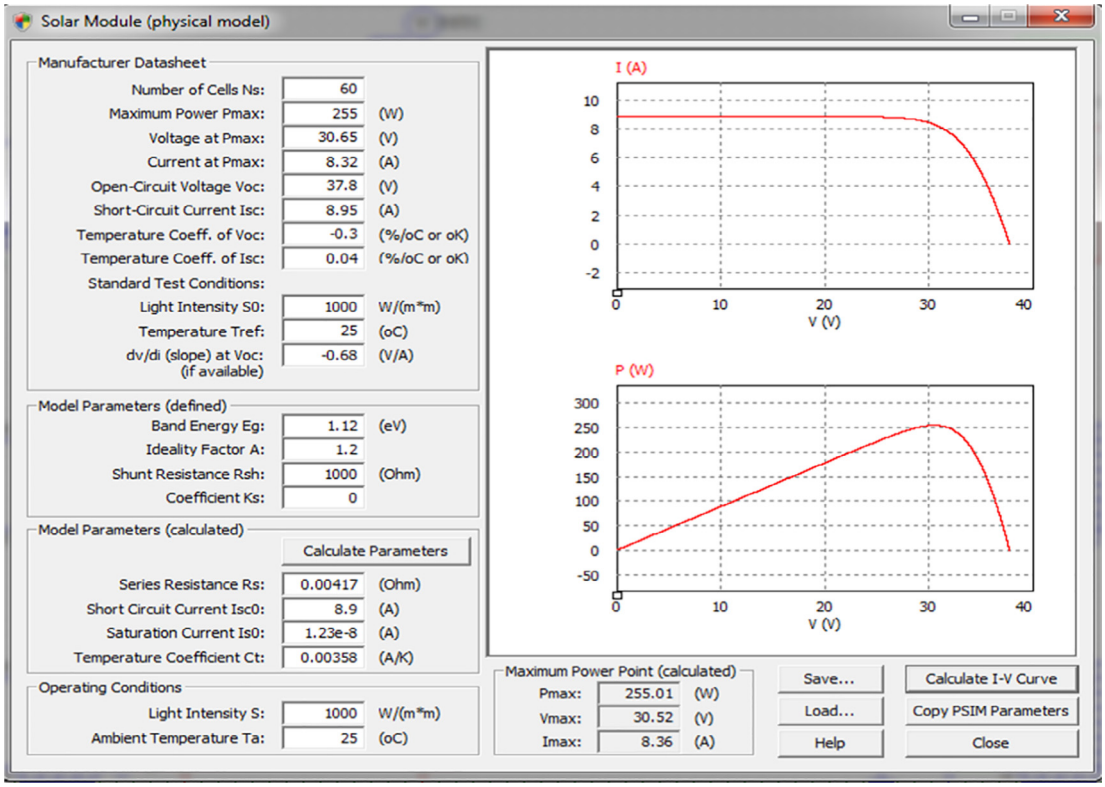

Fig. 1. The physical model of the PV array in PSIM.

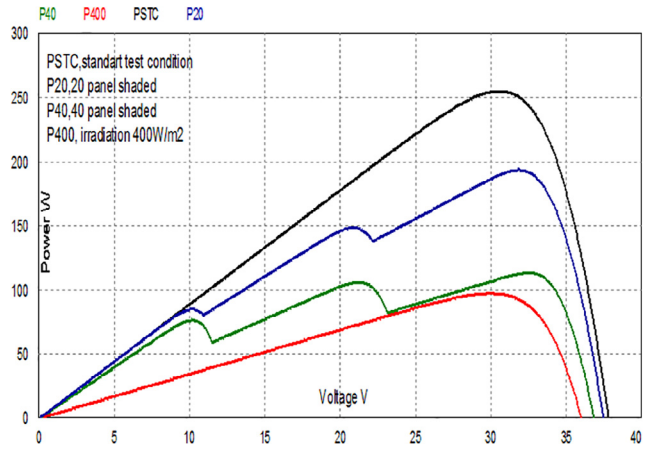

Fig. 3. Comparison of P-V curves obtained from different faults in PV array.

are modeled as partial shading faults, open circuit and short circuit faults.

4. Experimental results and discussions

4.1. Collecting the fault data

The PV simulator is a repeatable replication of PV array system which enables testing of new algorithms and estimation of system efficiency under different environmental conditions. The circuit simulation tools provide flexible, convenient, and economical ways for testing the PV system. PSIM simulation tools are promising circuit simulation softwares due to their fast simulation speed, immunity to convergence problem, library extensibility using C-code block, and achievability results compatible with the experimental data. When the simulation results obtained from PV Array modeling were compared with the laboratory results, it is seen that the obtained data are very close to the experimental data [40–41].

In the experimental study, a 3 × 4 PV power system is created by connecting 4 arrays in series and 3 arrays in parallel. In the case of STC, the output power obtained from the system is 3060 W, the open-circuit voltage is 151.2 V, and the short-circuit current is 26.85A. The PV system parameters used in the experimental study are given in Table 2.

The I-V and P-V curves for different irradiance values are shown in Fig. 4a and Fig. 4b. It can be seen in Fig. 4 that when the irradiation value decreases and the temperature value constants at 25 °C, the PV array short-circuit current changes. However, when the irradiation increases from 300 to 1000, the voltage only increases by 1 V. Therefore, the reduction in irradiation greatly changes the short-circuit current of the PV array.

Fig. 5a and Fig. 5b show the I-V and P-V curves obtained when the temperature value changes. As shown in Fig. 5, generally, when the solar irradiation is constant at 1000, if the PV array temperature rises, the open-circuit voltage decreases and the short-circuit current increases very little. Therefore, the open-circuit voltage of the PV array changes too much when the temperature changes.

Table 2
Parameters of PV system using for the experimental study.

Parameters	Value
Maximum Power, P_{max}	3060 W
Current at Pmax, I_{mpp}	24.96A
Voltage at Pmax, V_{mpp}	122.6 V
Short-circuit Current, I_{sc}	26.85 A
Open-circuit voltage, V_{oc}	151.2 V

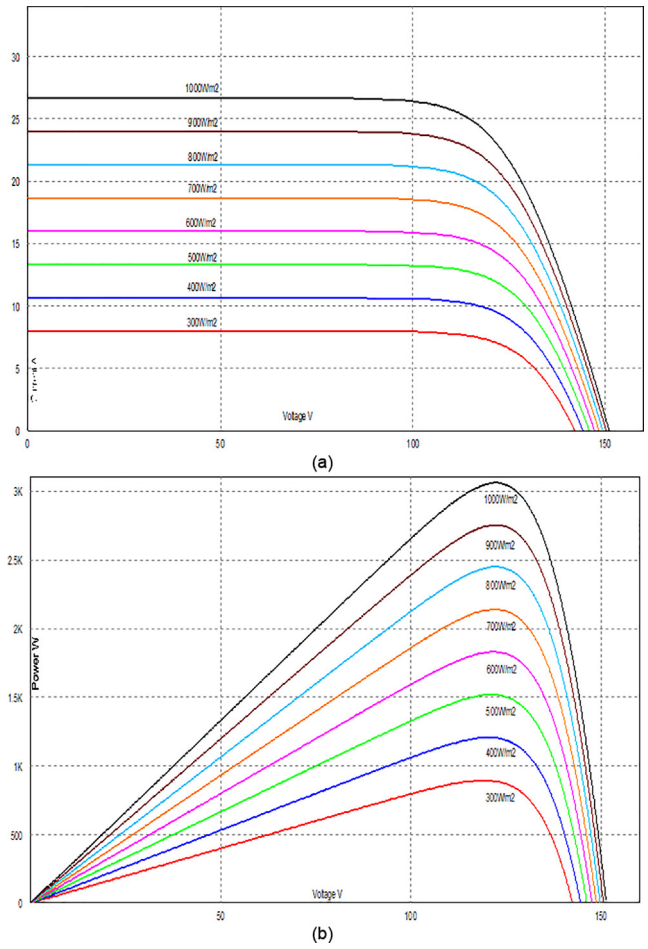

Fig. 4. (a) I–V curves of (b) P–V curves obtained from irradiation change at PV array.

When the PV power system is operating under standard test conditions, if a fault occurs in the PV system, the obtained I-V curves differ according to the type of fault. In the first experimental study, the open-circuit fault, short-circuit fault, partial shading fault are performed. The four different I-V curves obtained from the standard test condition, open-circuit fault, short-circuit fault, partial shading fault in the PV power system are given in Fig. 6a. As it can be seen in Fig. 6b, there are significant differences in the maximum output power of the PV array in the case of a fault.

In the second experimental study, the series resistance degradation faults are performed for three values of series resistance ($Rs = 0.00417\Omega$) $Rs + 50\%Rs = 1.5Rs = 0.00625\Omega$, $Rs + Rs = 2Rs = 0.00834\Omega$ and $Rs + 2Rs = 3Rs = 0.01251\Omega$. The four different I-V curves obtained in the standard test condition, three series resistance degradation fault in the PV power system are given in Fig. 7. Since the series resistance value is very small, it may be neglected in some cases. However, since it has an effect on the output power of the PV array and open-circuit voltage, it is important to determine the resistance faults. Reductions in series resistance cause an open-circuit voltage Voc decrease in the I-V curve. Fig. 8 shows the P-V curves for series resistance degradation fault. As shown in Fig. 8, the change of the series resistance results in the deviation of the maximum power point.

4.2. PV Fault detection and classification

Two different experimental studies for two different types of PV fault diagnosis are performed in this study. One type includes the open-circuit fault, short-circuit fault, and partial shading fault

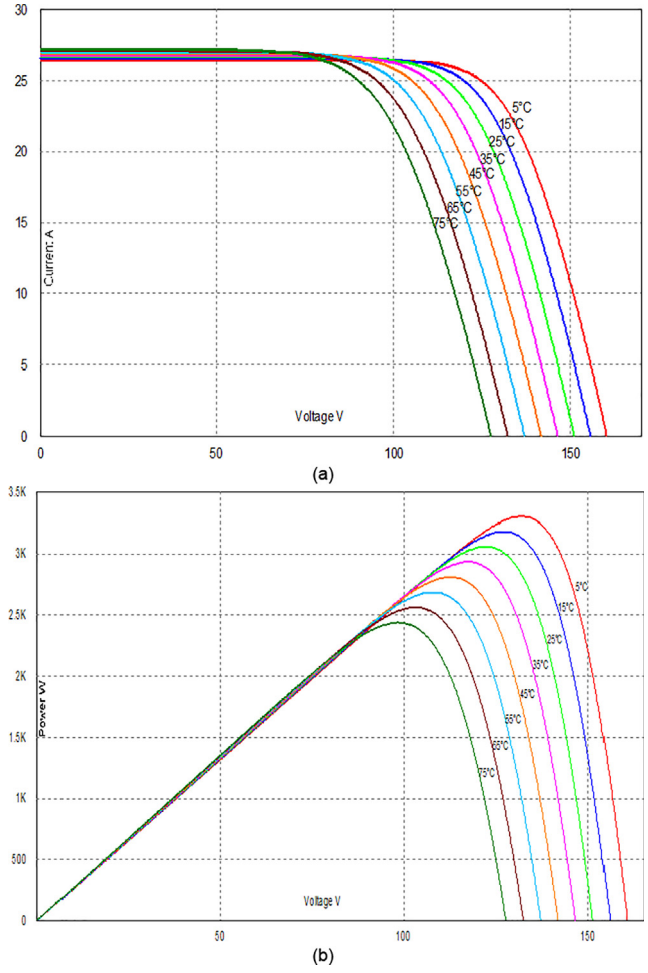

Fig. 5. (a) I–V curves of (b) P–V curves obtained from temperature change at PV array.

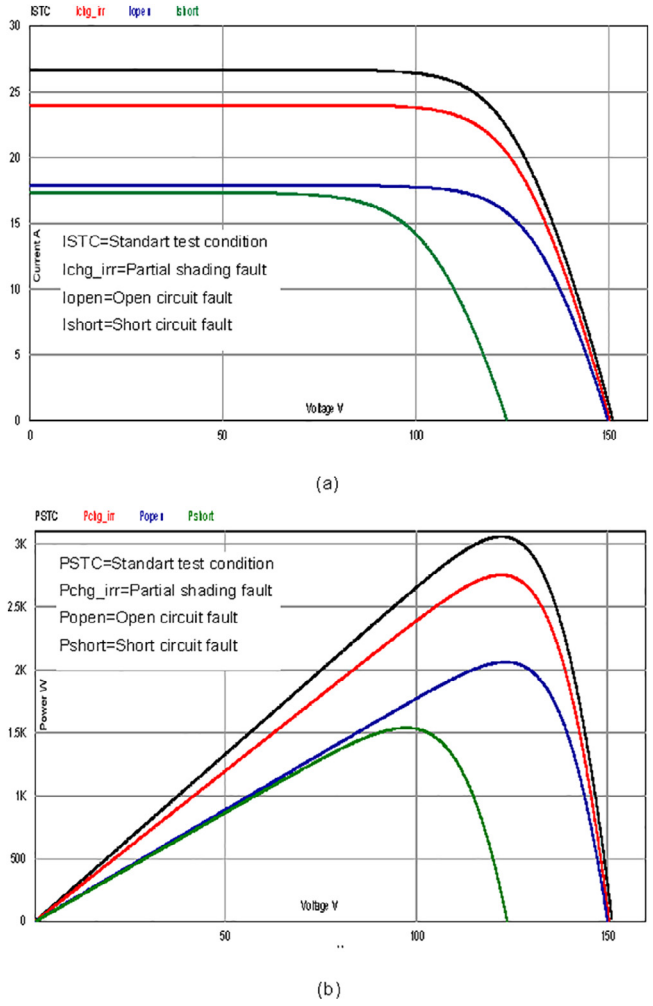

Fig. 6. (a) I–V curves, (b) P–V curves for different PV array fault.

while the other one degradation of series resistance. PV faults for the application are simulated using the PSIM program. The data obtained by the simulation in PSIM are analyzed with the DCVA method proposed in this study and in order to evaluate the obtained results, they are analyzed using the PCA method, too.

Experimental Study 1: When the I-V curves obtained in different irradiation values, different temperature values and different fault conditions are examined in the PV power system consisting of 3×4 PV array, it is seen that the I-V curves contain the information related to the type of fault occurring in the PV arrays. Short-circuit current and open-circuit voltages of PV array vary in case of fault. For the output power changes in the faults that can occur in PV arrays, the short-circuit current, the output power and the open-circuit voltage can be selected as the input parameters of the diagnosis and classification algorithm in array faults. The cause of using DC parameters in the experimental studies is the fact that these values are obtained only from the output of the PV array. DC parameters of a array are composed from short-circuit current, open-circuit voltage, current in maximum-power, voltage in maximum-power, and maximum power. If a fault condition occurs on PV systems, these parameters change. I-V and P-V curves corresponding to three different faulty conditions and the standard test condition are given comparatively in Fig. 6 and all the data (X) that are obtained from the simulation has a form as:

$$X = [V \quad I \quad P]_{1x3} \tag{26}$$

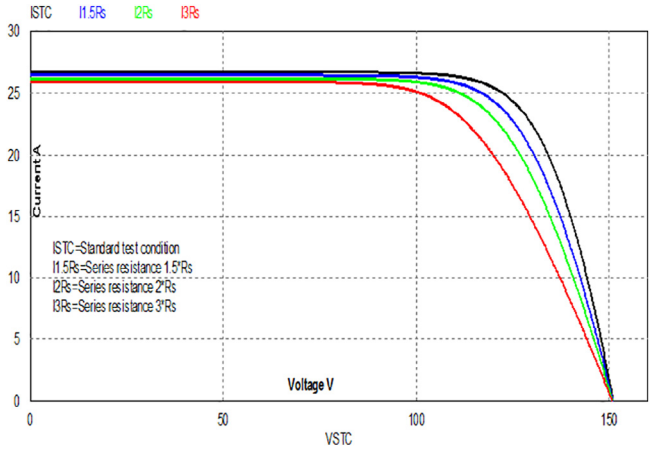

Fig. 7. The I-V curves for series resistance degradation fault.

The $1x3$ dimensional simulation data X are composed of voltage, current and power values of four data classes that each class corresponds to a different operation condition (class, $C_i, i = 1, .., 4$). These operation conditions include one healthy condition (STC), C_1, and three faulty conditions. These faulty conditions are partial shading, C_2, open-circuit, C_3, and short circuit, C_4, as given in Fig. 6. The fault diagnosis and classification config-

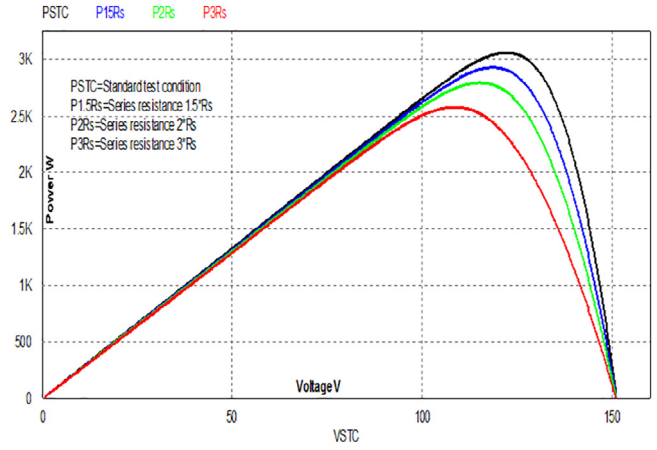

Fig. 8. The P-V curves for series resistance degradation fault.

uration constructed in this study given in Fig. 6 is just like the configuration that is constructed given in reference [16]. In reference [16], they proposed to use a statistical dimension reduction method, PCA that is popular in recent years, for PV fault detection and classification. However, PCA is a dimension reduction method rather than a classification method so its classification performance is poor according to other classification methods. Based on this fact, in this study, we proposed to use DCVA which has a high classification performance instead of PCA for the PV fault diagnosis system. In order to evaluate the fault detection and classification performance of the novel proposed method, DCVA in this study, the simulation data is also classified using PCA. The total simulation data used in this study is given in Fig. 9.

In two of the experimental studies, 10 fold cross-validation is used for model performance evaluations. That's why the total simulated data given in the form as in Fig. 9 is divided into ten parts. 10 model performance evaluations are performed in which each part is used one time as a training set, and one time as a test set and, the average of these 10 evaluations is taken as the model performance.

In the model construction step, a Discriminative Common Vector (DCV) for each of the four classes corresponding to one of the different operation conditions such as STC, partial shading, open

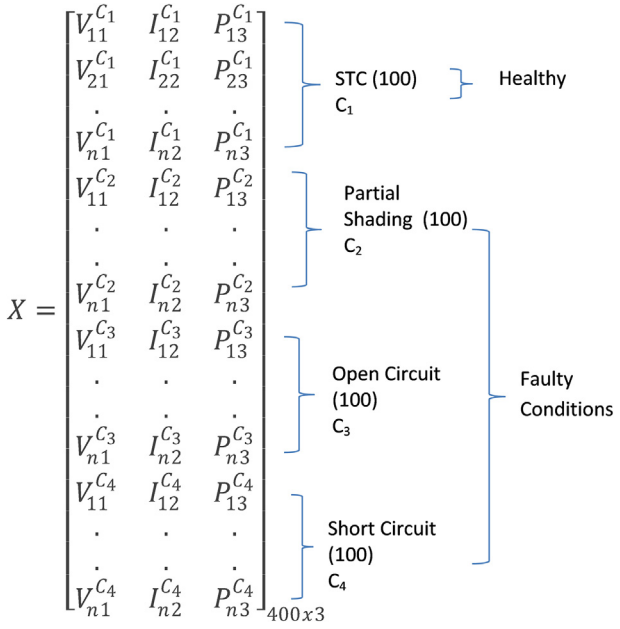

Fig. 9. The total simulation data used in the experimental studies.

circuit, or short circuit is obtained and used for classification purposes. The model construction steps for the proposed PV fault detection and classification method in this paper are as follows:

Step 1: A common matrix x^i_{com} is found for each operation condition. For this, the within-class scatter matrix S_W with the dimension 3x3 is obtained using Eq.(3). Applying Eigen decomposition on this scatter matrix three eigenvalues and corresponding three eigenvectors are obtained. To obtain x^i_{com} for each class, only the eigenvectors corresponding to the eigenvalues that are close to zero are used. As explained in Eq.(18) common matrices are found for each of the four classes. The feature space represented with the eigenvectors corresponding to the eigenvalues close to zero is the space that contains the common matrices for all classes. The flow chart of Step 1 is as given in Fig. 10.

Step 2: In this step, total scatter of the common matrices, *S_com* is obtained using Eq.(9) and then, Eigen decomposition is applied to this scatter matrix. Then, the Discriminative Common Feature

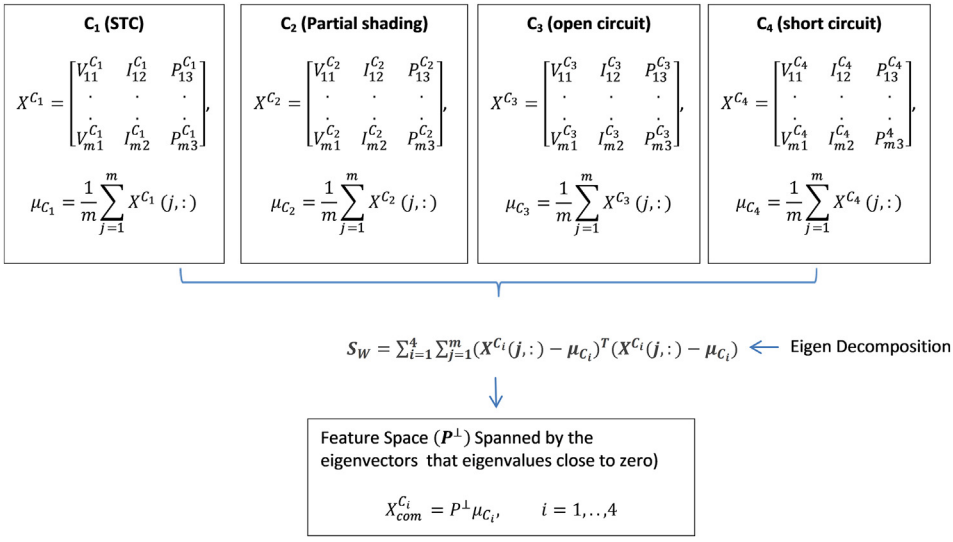

Fig. 10. The flow chart of Step 1.

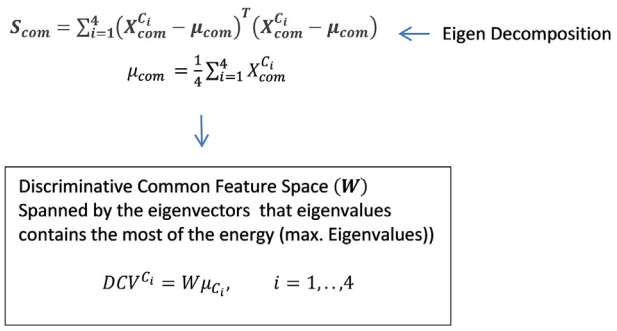

Fig. 11. The flow chart of Step 2.

space is obtained using the eigenvectors that maximize this scatter matrix and by projecting the class mean vector onto the Discriminative common feature space, a DCV is obtained for each class. The flow chart of Step 2 is as given in Fig. 11.

The pseudocode of DCVA that is used in this study to detect and classify the faults is as given in Table 3.

The model performance evaluation step for the proposed PV fault detection and classification method in this paper is as follows: In this step, the performance of the proposed method is evaluated using the separated simulation data. In this step, data that will be classified is projected onto the Discriminative Feature Space using the eigenvectors obtained in Step 2 of the model construction step. This projection vector is called the remaining vector. Then, Euclidian Distances between the remaining vector and the DCV vector of each class are computed. The data to be classified will be assigned to the class for which the distance is minimum, given as in Table 4.

The *k*-fold cross-validation method is used for evaluating the system performance. In fact, cross-validation is a statistical resampling method used to evaluate the classification performance on data that the classification model does not see, as objectively and accurately as possible. In this classification model, the collected data is divided into two in order to be used for creating a model and evaluating the model. However, since different accuracy values are obtained as a result of different divisions with this method, the performance of the model cannot be evaluated objectively. One of the most basic cross-validation methods is *k*-fold cross-validation. In this method, first, the data set is divided into "*k*" groups. Second, *k* piece classification studies (experimental studies) are performed as such the each of the *k* groups will be used for model construction once, and the remaining groups as data samples for the model performance evaluation. Then, the average of the total *k* piece experimental study results is taken as the model performance. In the literature, usually, 10 is chosen as the " *k* " value.

Table 4
Classification process of a new data sample.

Class DCVs	Euclid Distances $DCV_{remain} = WX_{classified}$	Assigned Class Label
$DCV^{C_1} = W\mu_{C_1}$	$< DCV^{C_1}, DCV_{remain} >^{1/2}$	Data will be classified is assigned to the class that the Euclid Distance is minimum
$DCV^{C_1} = W\mu_{C_2}$	$< DCV^{C_2}, DCV_{remain} >^{1/2}$	
$DCV^{C_3} = W\mu_{C_3}$	$< DCV^{C_3}, DCV_{remain} >^{1/2}$	
$DCV^{C_4} = W\mu_{C_4}$	$< DCV^{C_4}, DCV_{remain} >^{1/2}$	

In Fig. 12, the data projected onto the PCA feature space for the first experimental study is given. In this feature space, it can be seen that although the standard test condition and the partial shading fault can be separated accurately, the open-circuit and the short-circuit faults cannot be separated from each other.

On the contrary, in the same fault conditions, using the DCVA method which is derivated for the sufficient data case in this study as a first in the literature, we obtained extremely high classification rates. The reason can easily be seen from the DCVA feature space for the first experimental study given in Fig. 13. In DCVA feature space, scatter lines of the projected data for standard test condition, open-circuit fault, short-circuit fault and partial shading fault can easily be separated from each others.

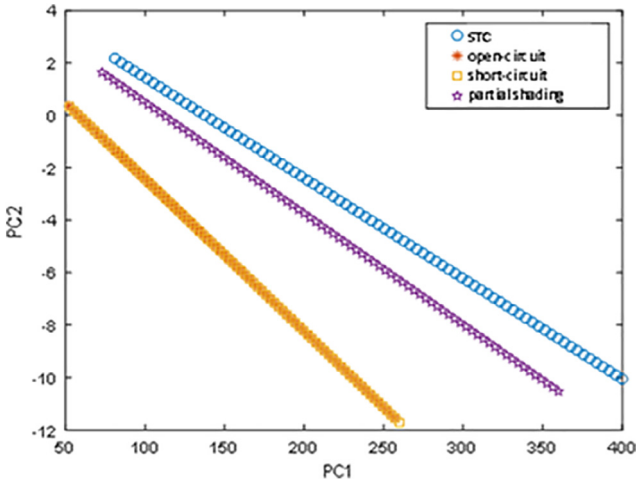

Fig. 12. Data projected onto the PCA feature space for the first experimental study.

Table 3
Puseudocode for fault diagnosis system.

This algorithm obtains Discriminative Common Vectors for sufficient data case.
Data samples: $x_j^i = [\mathrm{V} \quad \mathrm{I} \quad \mathrm{P}]_{1x3}, j = 1,2,..,m, i = 1,2,..,C$
Step1: Common vector for each class
Compute S_w: Within-class scatter matrix of all classes as in Eq. (3).
Perform Eigen analyse of S_w where $\mu_{C_i}, i = 1,..,4$ are the class mean vectors and Feature Space ($\boldsymbol{P}^{\perp}$) spanned by the eigenvectors (column vectors of matrix Q) that eigenvalues close to zero.
Common vector for each class:
for $\boldsymbol{i} = 1 : \boldsymbol{C}$
$X_{com}^{C_i} = P^{\perp}\mu_{C_i}, \quad \mathrm{P}^{\perp} = \sum_{j=1}^{k} \overline{\mathrm{QQ}}^{\mathrm{T}} \quad i = 1,..,4$**end**
Step2: Discriminative common vector for each class
Compute S_{com}: Within-class scatter matrix of common vectors of all classes as in Eq. (9). Perform Eigen analyse of S_{com} where $\mu_{C_i}, i = 1,..,4$ are the class mean vectors and Discriminative Common Feature Space ($\boldsymbol{W}$) spanned by the eigenvectors (column vectors of matrix $\boldsymbol{W}$) that contains the 95% of the energy.
Discriminative Common Vector for each class:
for $\boldsymbol{i} = 1 : \boldsymbol{C}$
$DCV^{C_i} = W\mu_{C_i}, \quad i = 1,..,4$**end**

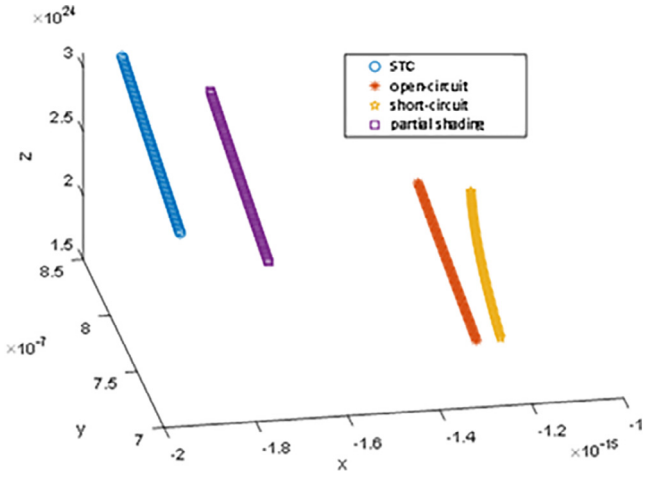

Fig. 13. Data projected onto the DCV feature space for the first experimental study.

Table 5
Classification percentages.

Method	DCVA	PCA
Classification performance (%)	99	95

These classification percentages obtained from the first experimental study performed using parameters of PV system are given in Table 2, and Table 5. At the same time, by balancing the number of samples for each class, the use of metrics such as the accuracy, for the performance evaluation has a high significance. Thus, the confusion matrix for the first experimental study both for the proposed method (DCVA) and also PCA is shown in Fig. 14 a and b, respectively.

Diagonal cells in the confusion matrix correspond to correctly classified observations. Cells outside the diagonal correspond to incorrectly classified observations. Each cell displays both the number of observations and the percentage of the total number of observations. The cell in the bottom right of the matrix shows the overall accuracy, which is also known as the accuracy metric. The overall accuracy is defined as the rate of the total correctly classified fault samples to the total sample number. According to the results obtained from the first experimental study, the proposed DCVA method provided significantly better classification rates than PCA. The classification performances are obtained 99% for DCVA and 95% for PCA method as can be seen from the classification rate given in Table 5 and the accuracy metric obtained from the confusion matrix given in Fig. 14. This is because PCA is a dimension reduction method and it eliminates the useful information required to discriminate the different classes from each other. But DCVA method keeps the specific and distinctive information for each class. Therefore, DCVA is more successful than PCA in classification.

Experimental Study 2: In the second experimental study, a series resistance degradation fault diagnosis with DCVA is carried out. In this fault condition, in order to evaluate the performance of the DCVA method, the data is also classified using PCA. The data is constructed from the simulation results of STC and three series resistance degradation fault. The 1x3 dimensional simulation data X composed of voltage, current, and power includes four data classes that each class (C1, C2, C3, C4) corresponds to a different operation condition. These operation conditions are one healthy condition (STC), C1, and three serial resistance degradation faults such as $R_{sadd} = \%50Rs$, C2, $R_{sadd} = Rs$, C3, and $R_{sadd}2Rs$, C4, as given in

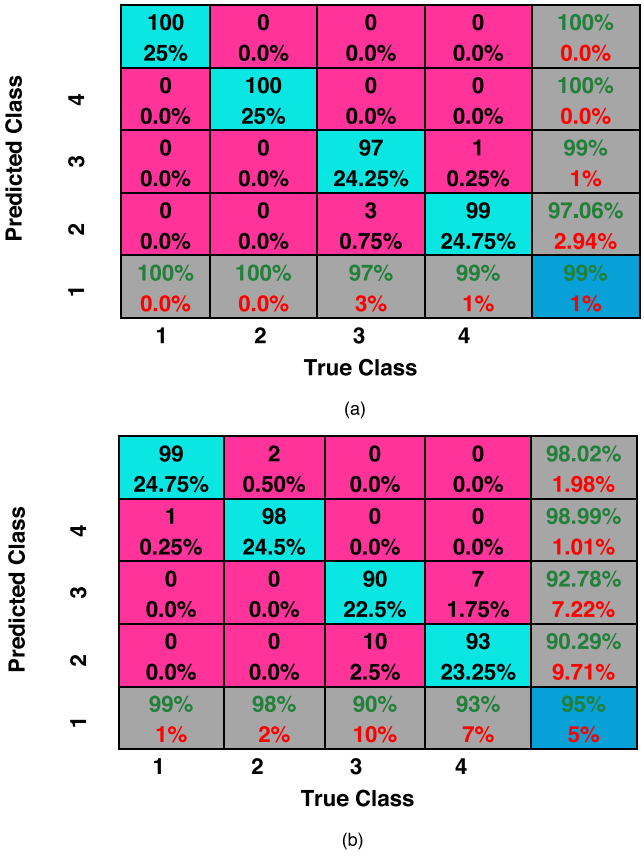

Fig. 14. Confusion matrix of the first experimental study: a) For the proposed method (DCVA), b) For PCA.

Fig. 7. The fault diagnosis and classification configuration given in Fig. 7 is just like the one in the first experimental study.

In the second experimental study, in the model construction step, a Discriminative Common Vector (DCV) for each class which corresponds to STC and three series resistance degradation fault is obtained and used for classification purposes. 10 fold cross-validation is used for model performance evaluation as it is in the first experimental study.

In Fig. 15, the data projected onto the PCA feature space for the second experimental study is given. As can be seen in Fig. 15, using

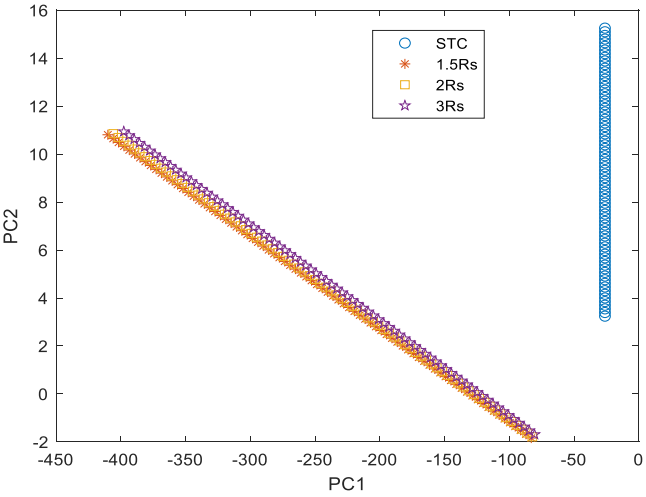

Fig. 15. Data projected onto the PCA feature space for the second experimental study.

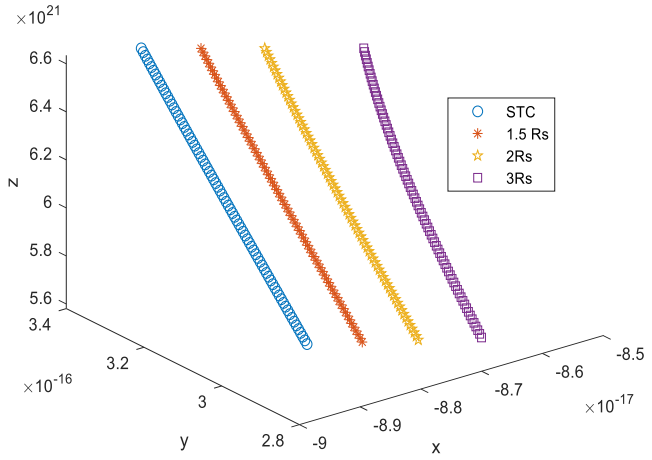

Fig. 16. Data projected onto the DCV feature space for the second experimental study.

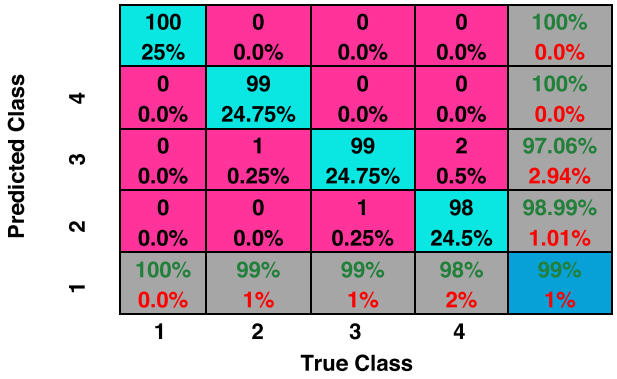

Fig. 17. Confusion matrix of the second experimental study for DCVA.

PCA we cannot classify the Rs degradation faults. Because all the different classes, healthy, $Rsadd = \%50Rs$, $Rsadd = Rs$, and $Rsadd = 2Rs$ are projected on very close to each other in the PCA feature subspace. That's why the different classes cannot be differentiated. Degraded series resistance fault diagnosis is studied using PCA in the literature. They had also obtained the same results. They stated that the additional data processing or/and additional information should be done or included in their study in order to separate the healthy case from the degraded series resistance [28].

On the contrary, in the same fault conditions, using the DCVA method which is derivated for the sufficient data case in this study as a first in the literature, we obtained an extremely high classification rate that is %99. The reason can easily be seen from the DCVA feature space for the second experimental study given in Fig. 16. In DCVA feature space, scatter lines of the projected data for STC and three series resistance degradation fault can easily be separated from each other. The confusion matrix of the second experimental study is also given in Fig. 17 and the accuracy metric obtained from the confusion matrix for this experiment study is 99%.

5. Conclusions

DCVA is a statistical pattern recognition method and it is defined in case of insufficient data case. It uses the null space of the within-class scatter matrix of all classes to define a common vector for each class in its first stage. However, in case of sufficient data, it cannot be used as the null space approaches zero. In this paper, a derivation of DCVA in case of sufficient data is proposed. The performance of the novel proposed derivation of DCVA is evaluated in a fault diagnosis of a PV system, in which the sample dimension is less than the sample number, (sufficient data case). To the best of our knowledge, this usage of DCVA for PV diagnosis systems in literature is a first. To prove the accuracy of the proposed method, the fault diagnosis is also performed using PCA which is also a statistical pattern recognition algorithm and recently became popular in PV fault diagnosis.

The data including voltage, current, and power used in the study is simulated using PSIM. Two different fault diagnosis are performed in the study. In the first one, the fault conditions are chosen as to be open-circuit, short-circuit, and partial shading. Thus, a total of four simulations are carried out, including one standard test condition and three fault conditions. In the second one, the fault conditions are chosen as to be three series resistance degradation. Thus, a total of four simulations are carried out, including one standard test condition and three fault conditions. The usage of novel proposed derivation of DCVA in case of sufficient data, in fault detection and classification problem for a PV system provides extremely high performances. The cause can be seen from Fig. 13 and Fig. 16 in which the data projected onto the DCV feature space is shown and from Fig. 12 and Fig. 15 in which the data projected onto the PCA feature space is shown. As can be seen from Fig. 13 and Fig. 16, the classes could be differentiated from each other using DCV successfully in the first experimental study, however, as can be seen from Fig. 12 and Fig. 15 using PCA, fault classes cannot be discriminated correctly.

In this paper, the advantages of using DCVA instead of using PCA in PV arrays fault detection and classification are analyzed. According to the results obtained from the first experimental study, the proposed DCVA method provided significantly better classification rates than PCA. The classification performances are obtained 99% for DCVA and 95% for PCA method as can be seen from the classification rate given in Table 5 and the accuracy metric obtained from the confusion matrix is given in Fig. 14. As for the results obtained from the second experimental study, the proposed method achieved 99% classification performance, however, PCA couldn't discriminate the faults. The reason can be explained by the fact that the PCA method eliminates the zero eigenvalues of the covariance matrix of an attribute space in transition to feature subspace. But, this causes a loss of information required to separate the different classes. Thus, PCA is applied for the reduction in size rather than classification. However, since DCVA maintains the indifference subspace that is spanned by the eigenvectors corresponding to the zero eigenvalues for classes, it is used to discriminate classes. Because DCVA holds specific and distinctive information for each class while PCA does not hold [32].

Contributions of this paper can be summarized as follows: First, derivation of DCVA in case of sufficient data which is novel in the literature is proposed. Second, DCVA that keeps the specific and distinctive information for a sample class is used for the first time in a PV system fault diagnosis in the literature. Third, the proposed method for PV fault diagnosis only uses voltage, current and power values of a PV system, that its sensitivity to environmental factors is low. In addition, it performs the detection and the classification independently of any PV sizes. The cost of determining and classifying PV array fault does not increase with the algorithm proposed.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] P.C. Chen, P.Y. Chen, Y.H. Liu, J.H. Chen, Y.F. Luo, Sol. Energy 119 (2015) 261–267, https://doi.org/10.1016/j.solener.2015.07.006.
[2] Firth, S. K.; Lomas, K. J.; Rees, S. J. Sol. Energy 2010, 84(4), 624, doi:org/10.1016/j.solener.2009.08.004.
[3] Q. Zhao, S. Shao, L. Lu, X. Liu, H. Zhu, Energies 11 (1) (2018) 238, https://doi.org/10.3390/en11010238.
[4] M. Hejri, H. Mokhtari, IEEE J. Photovolt 7 (1) (2016) 250, https://doi.org/10.1109/JPHOTOV.2016.2617038.
[5] T. Takashima, J. Yamaguchi, M. Ishida, Prog. Photovolt. 16 (8) (2008) 669–677, https://doi.org/10.1002/pip.860.
[6] T. Takashima, J. Yamaguchi, K. Otani, T. Oozeki, K. Kato, M. Ishida, Sol. Energ. Mat. Sol. C. 93 (6–7) (2009) 1079–1082, https://doi.org/10.1016/j.solmat.2008.11.060.
[7] N. Gokmen, E. Karatepe, B. Celik, Silvestre, S, Sol. Energy 86 (11) (2012) 3364–3377, https://doi.org/10.1016/j.solener.2012.09.007.
[8] S.R. Madeti, S.N. Singh, Energy 134 (2017) 121–135, https://doi.org/10.1016/j.energy.2017.06.005.
[9] S.R. Madeti, S.N. Singh, Sol. Energy 157 (2017) 349–364, https://doi.org/10.1016/j.solener.2017.08.047.
[10] Wu, L., Chen; Z., Long, C.; Cheng, S.; Lin, P.; Chen, Y.; Chen, H. App. Energ., 2018, 232, 36-53. doi:10.1016/j.apenergy.2018.09.161
[11] S. Kaplanis, E. Kaplani, Simul. Model. Pract. Th. 19 (4) (2011) 1201–1211, https://doi.org/10.1016/j.simpat.2010.07.009.
[12] B.K. Kang, S.T. Kim, S.H. Bae, J.W. Park, IEEE T Energy Conver. 27 (4) (2012) 885–894, https://doi.org/10.1109/TEC.2012.2217144.
[13] Zhu, Y.; Shi, X.; Dan, Y.; Li, P.; Liu, W.; Wei, D.; Fu, C. Int. Proc. Chinese Soc. of Electr. Eng. 2012, 32(4), 42-48.
[14] K.V.S. Bharath, F. Blaabjerg, A. Haque, M.A. Khan, Energies 13 (12) (2020) 3144, https://doi.org/10.3390/en13123144.
[15] K.V.S. Bharath, A. Haque, M.A. Khan, Applications of Computing, Automation and Wireless Systems in Electrical Engineering 553 (2019) 649–661, https://doi.org/10.1007/978-981-13-6772-4_56.
[16] Wang, J.; Gao, D.; Zhu, S.; Wang, S.; Liu, H. Energ. Source., Part A. 2019, 1-17, doi:10.1080/15567036.2019.1671557.
[17] P. Ray, D.P. Mishra, Eng. Sci. Technol. 19 (3) (2016) 1368–1380, https://doi.org/10.1016/j.jestch.2016.04.001.
[18] A. Haque, K.V.S. Bharath, M.A. Khan, I. Khan, Z.A. Jaffery, Energy Sci. Eng. 7 (3) (2019) 622–644, https://doi.org/10.1002/ese3.255.
[19] V.S.B. Kurukuru, A. Haque, M.A. Khan, A.K. Tripathy, IEEE Int. Conf. Computer Inform. Science (2019) 1–6, https://doi.org/10.1109/ICCISci.2019.8716442.
[20] A. Qazi, H. Fayaz, A. Wadi, R.G. Raj, N.A. Rahim, W.A.J. Khan, Clean. Proc. 104 (2015) 1–12, https://doi.org/10.1016/j.jclepro.2015.04.041.
[21] A. Youssef, M. El-Telbany, A. Zekry, J Renew. Sustain. Ener. 78 (2017) 72–79, https://doi.org/10.1016/j.rser.2017.04.046.
[22] Y. Wu, Q. Lan, Y. Sun, IEEE. Int. Conf. ICMA (2009) 2581–2585, https://doi.org/10.1109/ICMA.2009.5246742.
[23] M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales, P. Mather, Energy 140 (2017) 276–290, https://doi.org/10.1016/j.energy.2017.08.102.
[24] A. Belaout, F. Krim, A. Mellit, IEEE. Int. Conf. ICMIC (2016) 144–149, https://doi.org/10.1109/ICMIC.2016.7804289.
[25] Chen, Z.; L. Wu; S. Cheng; P. Lin; Y. Wu; W. Lin; Appl. Energ., 2017, 204, 912–31, doi:10.1016/j.apenergy.2017.05.034.
[26] Y. Zhao, R. Ball, J. Mosesian, J.F. de Palma, B. Lehman, IEEE T., Power Electr. 30 (5) (2014) 2848–2858, https://doi.org/10.1109/TPEL.2014.2364203.
[27] E. Garoudja, F. Harrou, Y. Sun, K. Kara, A.S. Chouder, Sol. Energy 150 (2017) 485–499, https://doi.org/10.1016/j.solener.2017.04.043.
[28] S. Fadhel, C. Delpha, D. Diallo, I. Bahri, A. Migan, M. Trabelsi, M.F. Mimouni, Sol. Energy 179 (2019) 1–10, https://doi.org/10.1016/j.solener.2018.12.048.
[29] L.C. Chen, P. Lin, J. Zhang, Z.C. Chen, Y.H. Lin, L.J. Wu, S.Y. Cheng, IOP C, Ser. Earth Env. 188 (1) (2018) 1–8, https://doi.org/10.1088/1755-1315/188/1/012089.
[30] Gulmezoglu, M. B.; Dzhafarov, V;. Edizkan, R.; Barkana A.; Comp. Speech Lang.2007, 21(2), 266-281, doi:10.1016/j.csl.2006.06.002.
[31] Cevikalp, H.; Neamtu, M;. Wilkes, M.; Barkana, A.; IEEE T. Pattern. Anal. 2005, 27(1), 4-13. doi: 10.1109/TPAMI.2005.9
[32] Gulmezoglu, M. B.; Dzhafarov, V.; Barkana, A; IEEE T. Speech Audi P. 2001, 9(6), 655-662. doi:10.1109/89.943343
[33] Morrison, D. F Multivariate Statistical Methods, McGraw-Hill, NewYork, 156–195, 1967.
[34] E. Oja, Subspace Methods of Pattern Recognition, John Wiley, & Sons Inc., New York, 1983.
[35] D.L. Swets, J. Weng, IEEE T. Pattern. Anal. 18 (1996) 831–836, https://doi.org/10.1109/34.531802.
[36] M.S. Ismail, M. Moghavvemi, T.M.I. Mahlia, Energ. Convers Manage. 73 (2013) 10–25, https://doi.org/10.1016/j.enconman.2013.03.033.
[37] D. Sera, R. Teodorescu, P. Rodriguez, IEEE Int. Symp. Ind. electron. (2007) 2392–2396.
[38] B. Sahoo, S.K. Routray, P.K. Rout, Eng. Sci. Technol. 21 (4) (2018) 639–653, https://doi.org/10.1016/j.jestch.2018.06.007.
[39] S. Motahhir, A. El Ghzizal, S. Sebti, A. Derouich, Int. J. Photoenergy (2018) 1–14, https://doi.org/10.1155/2018/3286479.
[40] J.Y. Park, S.J. Choi, Sol. Energy 90–98 (2017,145,), https://doi.org/10.1016/j.solener.2016.12.003.
[41] A. Salman, A. Williams, H. Amjad, M.K.L. Bhatti, M. Saad, *IEEE* Global Humanitarian Techn. Conf. (2015) 360–364, https://doi.org/10.1109/GHTC.2015.7343997.

Yasemin Onal, is an assistant professor at Electrical and Electronics Department of Engineering Faculty of Bilecik Seyh Edebali University (BSEU). She received her BS and MS degrees in Electrics and electronics from the University of Eskisehir Anadolu in 2011. Her current research interests include power electronic, signal processing, and power quality.

Umit Cigdem Turhal is an assistant professor at Electrical and Electronics Department of Engineering Faculty of Bilecik Seyh Edebali University (BSEU). She received her BS and MS degrees in Electrics and electronics from the University of Eskisehir Osmangazi in 2002 and 2008, respectively. Her current research interests include signal and image processing, data mining, machine learning and pattern recognition.