Photovoltaic Fault Detection and Classification Using Common Vector Approach Based on I–V Curve

U.C. Turhal and Y. Onal*

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

Photovoltaic systems are subject to various faults. These faults result in significant degradation in photovoltaic system performances. These degradations result in a significant reduction in system efficiency as a result of increase in operating costs and times of exclusion of the photovoltaic system. Therefore, health monitoring and high accuracy error detection are extremely important for a photovoltaic system. In this study, fault detection and classification in photovoltaic systems are performed by using common vector approach which is first proposed in the literature for speech recognition and then is applied for face recognition. The faulty conditions discussed in the study are partial shading and series resistance degradation. In both fault detection situations, the common vector approach method is used to determine the type of fault using simulation data obtained under healthy operation and faulty operation conditions and very high fault detection rates are obtained. In order to evaluate the accuracy of the proposed method, the data obtained are also evaluated with the principal component analysis method, which is previously presented for photovoltaic system fault detection in the literature. According to the results obtained, principal component analysis method completely fails in case of serial resistance degradation fault. However, by using common vector approach method proposed in this study, a very high fault detection rate such as 97.5% can be obtained in serial resistance degradation fault. Likewise, in the case of partial shading, higher fault detection rate is achieved in common vector approach method (99.6%) compared to principal component analysis method (95.4%).

1. Introduction

The solar energy has been used more widely than fossil energy because it does not cause air pollution. It has no energy costs and is an inexhaustible source of energy. Photovoltaic (PV) systems which have serial and parallel connected modules are used to convert solar energy into electrical energy [1].

In PV systems it is extremely important to identify situations where the operating conditions change. These cases are called "faults", they are either system-based or environmental factors based [2, 3]. Because these types of conditions, which can be called as bad working conditions, can seriously affect the produced energy yield, system efficiency, as well as the security and reliability of the entire system in case of them not being detected and corrected [4–6].

In recent years, many studies have been proposed for the fault diagnosis of PV systems. These studies generally differ in terms of main features such as rapid fault detection, required input data, and the separability of different faults [3]. Faults can be mainly grouped into two categories as to be visual and thermal methods, and electrical methods [7]. Electrical methods use electrical signals data as the input data to the fault diagnosis systems.

These methods can be classified among themselves as to be I–V characteristics analysis, power losses analysis, voltage and current measurement, statistical and signal processing approaches, and artificial intelligence techniques [8–10].

In this paper, a fault detection method that is based on artificial intelligence techniques, which belongs to electrical methods is proposed. In this study, to analyze thus effectively detect the faults of PV array outputs the common vector approach (CVA), that is a feature extraction method, is used [11]. To our best knowledge, CVA method has not yet been used for PV systems diagnosis in such operating conditions in the literature. In recent years, the use of the principal component analysis (PCA) method for detecting PV system faults has been frequently seen in the literature [12, 13]. The PCA, an optimal linear dimension reduction technique, in which data is projected into a lower dimensional feature subspace, is a feature extraction method like CVA method. There are many applications of PCA method in different diagnosis applications in the literature [14–17].

In this paper, the advantages of using CVA instead of using PCA in PV system fault detection and classification are analyzed. According to the experimental results CVA method is much superior over the PCA method in PV system fault detection and classification. Especially in series resistance degradation fault, while PCA cannot separate the faulty conditions, CVA provides an extremely high separability. PCA eliminates the features

————
*corresponding author; e-mail: yasemin.onal@bilecik.edu.tr

corresponding to the zero eigenvalues of the covariance matrix of an attribute space in the transition to a feature subspace, which also causes loss of information, required to separate the different classes. Therefore, it is applied for reduction in size rather than classification. However, since CVA maintains the indifference subspace for classes, it is used to discriminate classes because it holds specific and distinctive information for each class [18, 19].

In this study, the current, voltage, and power data are generated by using PSIM. PV panel configurations used in these simulations are generated as given in [12] in which fault detection is performed using PCA. According to the results of the analysis, it is seen that the CVA method, which is proposed for the first time in the literature by us in the PV system fault detection, gives considerably better results than the PCA method. As fault, two cases are handled that are given. The first one is a series of resistance degradation and the second one is a partial shading fault. Simulation results for the detection and classification of these faults are obtained under healthy conditions, series resistance degradation conditions, and partial shading conditions. According to the obtained results, PCA method completely fails in case of serial resistance fault. A very high fault detection rate such as 97.5% using CVA method is obtained in serial resistance fault. Likewise, in the case of partial shading, higher fault detection rate is achieved in CVA method (99.6%) compared to PCA method (95.4%).

This paper is organized as follows. The PV panel faults modeling in PSIM and the faults detection and classification using PCA and CVA methods are addressed in Sect. 2. The experimental results for the series resistance fault and partial shading fault using CVA and PCA are given in Sect. 3. Finally, brief conclusions are drawn in Sect. 4.

2. Modeling of PV panel faults

2.1. Mathematical model of PV panel

One of the most widely used models to reproduce the I–V characteristic of a PV panel is one or two diodes equivalent electrical model. To obtain the basic characteristics of the PV cell generally the classical single-diode electrical model is used. In this model, the PV panel is modeled as an electrical equivalent circuit using a very small series resistor R_s, a large shunt resistor R_{sh}, a parallel diode D. The photon current I_{ph} is the current generated proportionally by the surface temperature and insolation, and is shown as the photon current source in the equivalent circuit [20]. Figure 1 shows the electrical equivalent circuit of a PV cell. Other models have been developed to better model of physical events in a PV panel [21].
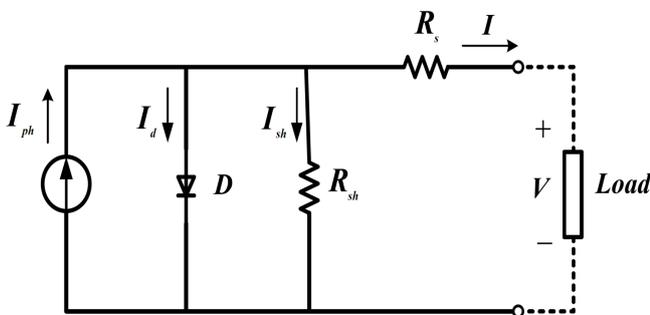
The output current I expressed as

$$I\left(1 + \frac{R_s}{R_{sh}}\right) = n_p I_{ph} - \frac{V - n_s}{R_{sh}}$$

$$-n_p I_{sat}\left[\exp\left(\frac{q\left[\frac{V}{n_s} + R_s I\right]}{A k_B T}\right) - 1\right] \qquad (1)$$

can be calculated using PV output voltage V, electron charge $q = 1.602176565 \times 10^{-19}$, Boltzmann's constant $k_B = 1.38 \times 10^{-23}$ J/K, PV's cell temperature T, number of serial cells n_s, and number of parallel cells n_p.

The cell saturation current I_{sat} expressed as

$$I_{sat} = I_{rr}\left(\frac{T}{T_r}\right)^3 \exp\left(\frac{qE_g}{A k_B T}\left(\frac{T}{T_r} - 1\right)\right) \qquad (2)$$

can be calculated using cell temperature T, the band energy of each PV cell E_g (for crystalline silicon $E_g = 1.12$ eV), and the reference temperature of the PV cell T_r. In the case of the photon current I_{ph}, one can use expression

$$I_{ph} = \left[I_{sso} + k_i\left(T - T_r\right)\right] \times \frac{S_i}{1000}. \qquad (3)$$

The value of I_{ph} depends on short circuit current I_{sso}, short circuit current temperature coefficient k_i, input light intensity S_i [W/m^2] (solar irradiance), and reference input temperature T_r (25 °C).

2.2. PSIM model of PV panel

PSIM package program is used in the design and simulation of power electronics and motor control circuits. It provides fast and accurate results in simulating and controlling power electronics circuits, and also in analog and digital motor control.

The photovoltaic panel used in this study is the FL60-250MBP PV monocrystal panel, whose parameter is given in Table I [22]. In the simulation circuit of the PSIM, a solar module is used, which can be used for solar radiation and ambient temperature changes. The solar module is shown in Fig. 2. Many parameter inputs are required to the solar module. Some of these parameters can be obtained from the manufacturer data sheets, while other panel parameters such as photocurrent, diode saturation current, series and shunt resistors, and ideality factor must be obtained by means of trial and error through the physical model. The shunt resistance R_{sh} is provided randomly to PSIM and then PSIM automatically calculate series resistance R_s.

Fig. 1. The equivalent circuit of a PV cell.
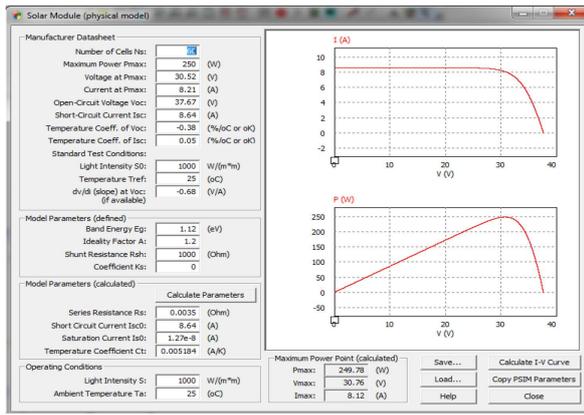
Fig. 2.   The physical model of the PV panel in PSIM.

TABLE I

PV panel physical model parameters.

number of cells	n_s	60
standard light intensity	S	$1000\ \mathrm{W/m^2}$
reference temperature	T_{ref}	$25\,^{\circ}\mathrm{C}$
series resistance	R_s	$0.00335\ \Omega$
shunt resistance	R_{sh}	$1000\ \Omega$
short circuit current	I_{sc0}	8.64×4 A
saturation current	I_{s0}	1.27×10^{-8} A
band energy	E_g	1.12
ideality factor	A	1.2
temperature coefficient	C_t	0.005184
coefficient	k_s	0
maximum power	P_{max}	250 W
current at P_{max}	I_{mpp}	8.21 A
voltage at P_{max}	V_{mpp}	30.52
short-circuit current	I_{sc}	8.64 A
open-circuit voltage	V_{oc}	37.67 V

The PV panel under test is composed of 60 PV cells connected in series. The simulation tests have been performed under several operating conditions:

- Series resistance degradation mode,

- Healthy condition mode,

- Partially shading modes with different shading applied on the PV cells.

Due to the activation of bypass diodes the faulty modes are represented differently in the I–V curve. In the following, we simulated a PV panel with the same specifications as the one considered in [12].

2.3. Modeling of healthy condition and series resistance degradation

Figure 3 shows the PSIM simulation circuit for the different series resistance degradation mode and the healthy condition mode. The solar irradiation is $1000\ \mathrm{W/m^2}$ and reference temperature is $25\,^{\circ}\mathrm{C}$ under the

Fig. 3.   The PSIM simulation circuit used for the series resistance degradation mode and healthy mode.
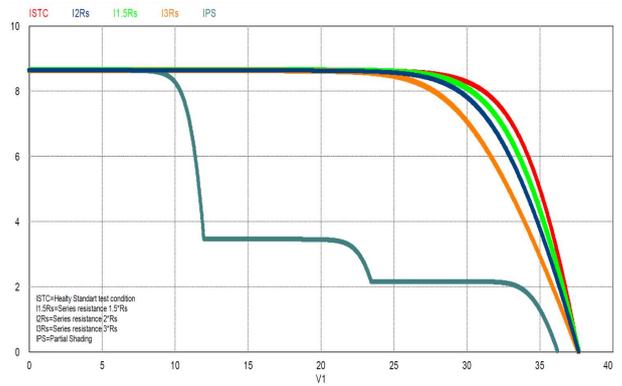
Fig. 4.   PV faults on the I–V curve.

standard test conditions (STC) in the simulation circuit. The simulation is made for four values of series resistance $(R_s = 0.0035\ \Omega,\quad R_s + 50\%R_s = 0.00525\ \Omega,$ $R_s + R_s = 0.007\ \Omega$ and $R_s + 2R_s = 0.0105\ \Omega)$. The I–V curves obtained from the simulation results of healthy mode, partial shading mode, and the series resistance degradation mode are shown in Fig. 4.

Since the series resistance value is very small, it may be neglected in some cases. However, since it has an effect on output power of PV panel and open-circuit voltage, it is important to determine the resistance degradation. Degradations in series resistance cause a open-circuit voltage V_{oc} decrease in the I–V curve. Figure 5 shows the P–V curves for different values of R_s. As shown in Fig. 5, the change of the series resistance results on the deviation of the maximum power point.

The I–V curves obtained from three different solar irradiation level when the PV panel operates in healthy mode are shown in Fig. 6. Solar irradiation are 850, 845, and $832\ \mathrm{W/m^2}$. The short-circuit current of the PV panel varies with solar radiation. The PV panel produces less current to low solar irradiation. Simulation results are consistent with the datasheet information given under STC with series resistance of $0.0035\ \Omega$.
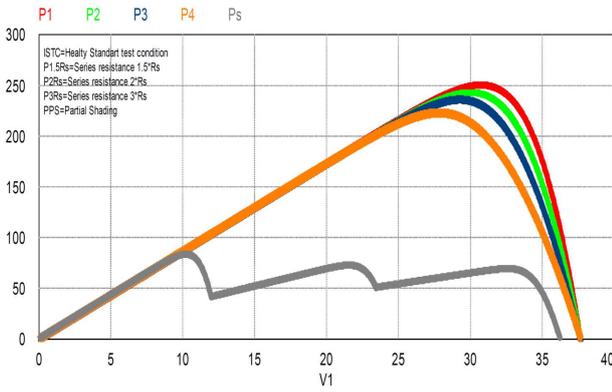
Fig. 5.   The P–V curves for different values of R_s.
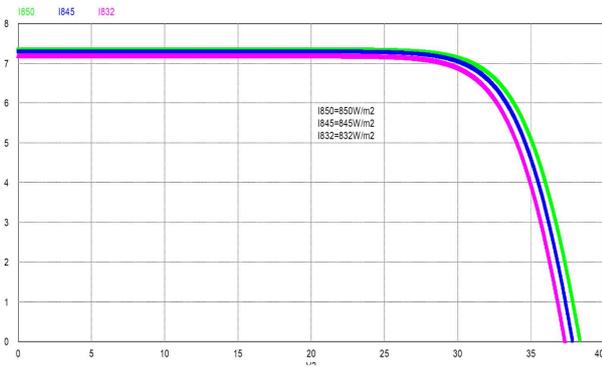
Fig. 6.   The I–V curves under healthy condition mode.

2.4. Analysis of the partial shading fault mode

Total or partial shading conditions cause very serious problems in PV systems. In PV panels, when the shaded cells of PV panel is less than the illuminated cell of the PV, the current produced by the illuminated cells is larger than the current produced by the shaded cells. The diode of shaded cells reverse biased and the power will be lost in the shaded cells. To overcome power loss, the PV panel is composed of solar cells connected by series and a bypass diode connected by parallel under the condition of partial shading.

The shading mode is applied to the three sub-strings as irregular irradiations. To simulate the effect of shading, the PV panel under test is composed of 60 PV cells connected in series and three sub-strings of 20 PV cells. Three sub-strings are equipped with three bypass diodes; each one is connected in antiparallel to protect a PV substring. Figure 7 shows the simulation circuit of the three sub-strings in PSIM.

Under partial shading, the current flows through the bypass diode instead of the shaded string for the bypass diodes are reversely biased. The power will not be lost in the shaded cells and only the illuminated cells generate power.
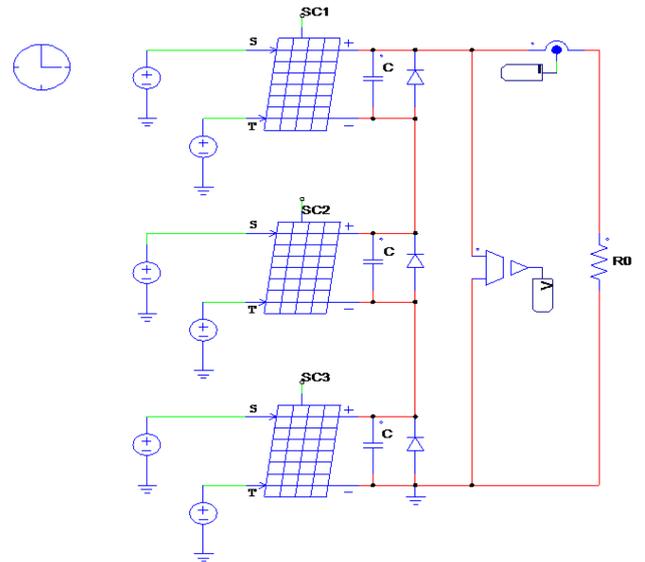
Fig. 7.   The simulation circuit of the three sub-strings for partial shading fault.
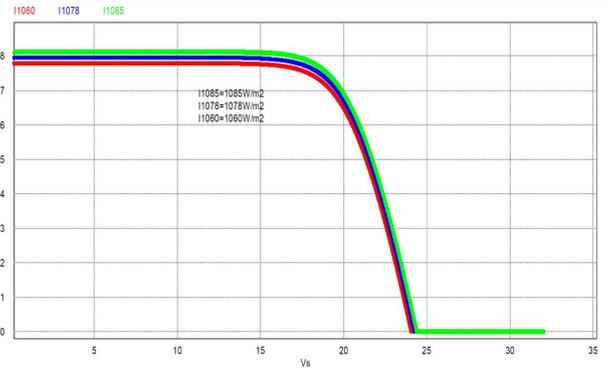
Fig. 8.   The I–V curves under 12 cell shading.

Figures 8–11 show the I–V curves for four configurations of shading. Multiple peaks occur in all I–V curves depending on the type of shading applied in the shaded conditions. These peaks indicate that the efficiency of the PV system is reduced because the maximum power is reduced under shading. The PV cells of the panel are tested under irregularly irradiation when partial shaded. Depending on the location and number of cells, in which partially shading occurs on the PV panel, the I–V curves obtained from the simulation results vary.

Figure 8 shows the obtained I–V curve when the first sub-string is partially shaded. The 12 PV cells of the first sub-string are shaded. The shaded cells close a bypass diode and the diode opens faulty sub-string. The closed diode prevents the shaded cells from heating up and becoming hot spots. Due to the loss of the substring, a steep drop in voltage occurs in the I–V curve. The 12 PV cells of the first sub-string and 6 cells of the second sub-string are shaded.
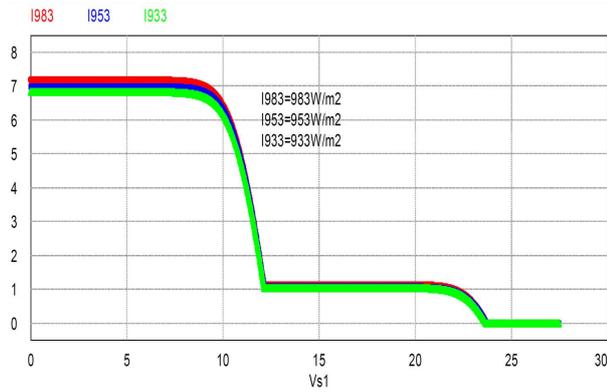
Fig. 9. The *I–V* curves under 18 cell shading.
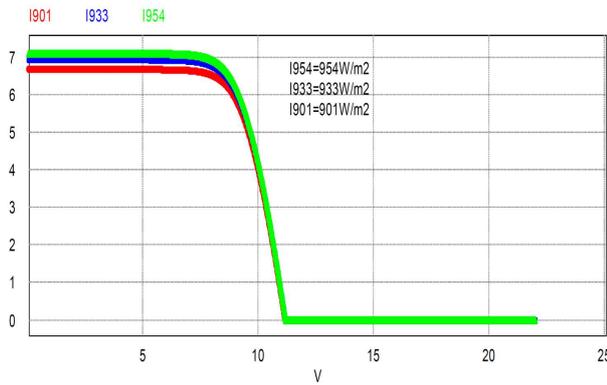
Fig. 10. The *I–V* curves under 24 cell shading.
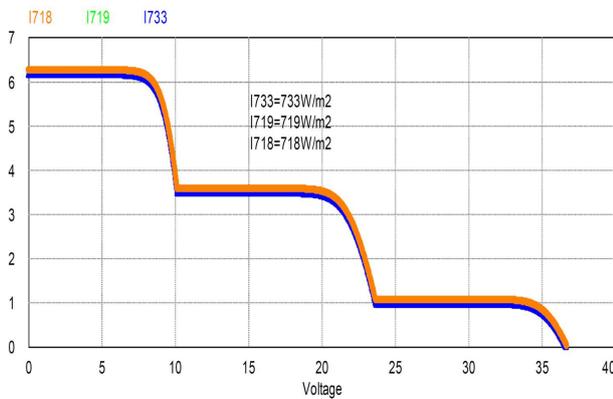
Fig. 11. The *I–V* curves under 2 cell shading.

Figure 9 shows the obtained *I–V* curve when the first and second sub-strings are partially shaded. The shaded cells close two bypass diode and the diodes open the faulty sub-strings. Figure 10 shows the 12 shaded cells of the first sub-string and 12 shaded cells of the second sub-string. The effect of shaded cells on the output current is clearly illustrated by the two peaks occurring on the *I–V* curve.

The one PV cells of the first sub-string and one PV cells of the third sub-string are shaded given as in Fig. 11. The effect of shaded cells on the output current is clearly illustrated by the three peaks occurring on the *I–V* curve.

2.5. Fault detection and classification

In this paper fault detection and classification in a PV system is performed using a supervised classification algorithm called CVA. CVA method can also be used both for fault detection and classification. In Fig. 4, the simulation results obtained under one healthy and four faulty conditions are given. Three of these faults are degradation of series resistances and one of these is partial shading.

The projection of the simulated data on CVA feature space is given in Fig. 12. As it can be seen from Fig. 12, CVA can detect and classify all the conditions given in Fig. 4 from each other successfully. However, the case is not the same for PCA. Using PCA only the partial shading fault detection can be performed but degraded series resistance cannot, as shown in Fig. 13.
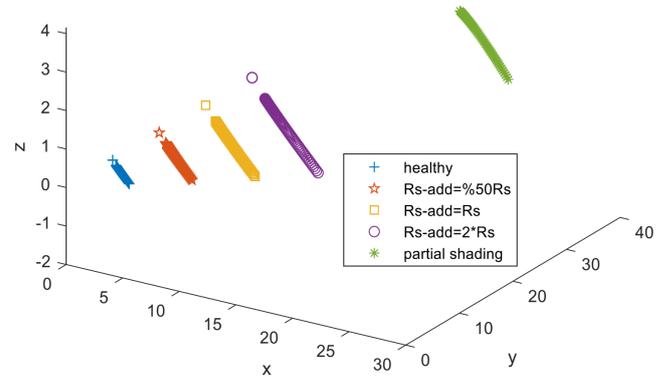
Fig. 12. Data projected onto CVA feature space.
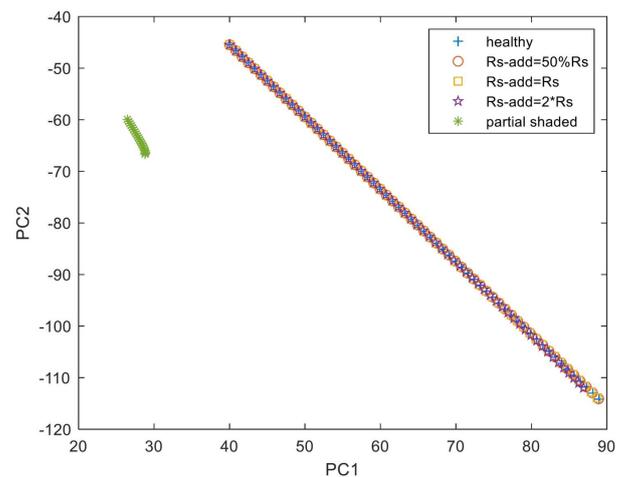
Fig. 13. Data projected onto PCA feature space.
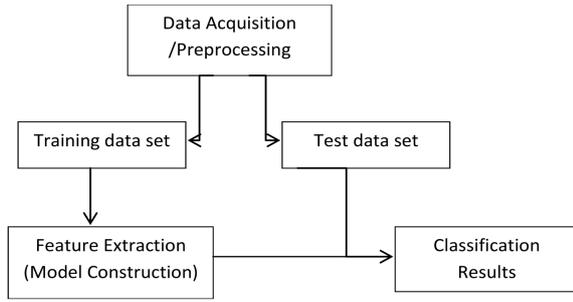
Fig. 14.   Supervised classification scheme.

A supervised classification scheme includes a training step and then a testing step as given in Fig. 14. That is why the dataset is first divided into two groups as one of them is for training and the other is for testing. In the training step, the classification model that practically corresponds to the feature extraction process, is constructed using the training data set. Once the model is constructed, then using the test data set classification performance of the model is evaluated. In this paper, the classification results are evaluated using the method PCA, for the same database for comparison purposes. Mathematical formulation of these two methods, CVA and PCA, are given below.

2.6. Feature extraction using PCA

The PCA method is the process of obtaining eigenvalues and eigenvectors of covariance matrices for normal distributions, also known as optimal linear dimension reduction [23, 24]. The aim of this method is the projection of the data from a high dimensional data space into a lower dimensional feature subspace. This projection is realized by using eigenvalues and eigenvectors obtained from eigenvalue-eigenvector decomposition of the covariance matrix of normal distribution. The PCA is based on the elimination of feature components along the eigenvectors corresponding to the smallest eigenvalues and keeping the feature components along the eigenvectors corresponding to the largest eigenvalues. Thus, this process reduces the number of dimensions in the feature space, however such projection often results in loss of information that distinguishes different classes.

All objects in the database are represented by vectors with the size of $1 \times m$. Here, m denotes the number of attributes of the objects. Thus, each object in the database corresponds to a point in a m-dimensional attribute space. With the PCA method applied in the attribute space of the database, the data is projected into a lower dimensional feature subspace, where correlation among them is reduced. The PCA gives the basis vectors for this feature subspace. Each basis vector is m-length and is the eigenvector of the covariance matrix with the size of $m \times m$, which is generated in the attribute space. The formulation of PCA is as given below.

Let us assume X_1, X_2, $\ldots, X_N$, where $\{X_i\} = (x_{i1}, \ldots, x_{im})$, is the training set of objects defined in the m-dimensional attribute space. The average of these objects $\bar{X}$ is understood as

$$\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i, \tag{4}$$

and the covariance matrix C can be obtained

$$C = \frac{1}{N} \sum_{i=1}^{N} (X_i - \bar{X})(X_i - \bar{X})^T. \tag{5}$$

Eigenvalue-eigenvector decomposition is applied to the covariance matrix given in (5). One can choose k significant eigenvectors $(e_1 e_2 \ldots e_k)$, which are corresponding to the k largest eigenvalues. These eigenvectors span the lower dimensional feature subspace. Taking the projection of an object X_i onto these eigenvectors, X_i can be defined in the feature subspace, as shown in (6):

$$Y_i = (X_i E)\, E, \quad i = 1, 2, \ldots, N,$$

$$E = \left\{ \begin{matrix} \uparrow & & \uparrow \\ e_1 & \ldots & e_k \\ \downarrow & & \downarrow \end{matrix} \right\}, \tag{6}$$

where Y_i is defined as the projection of the vector X_i, that belongs to the attribute space onto the feature subspace, and E is the eigenmatrix whose columns are constructed from the k largest eigenvalues.

2.7. Feature extraction using CVA

The CVA is a subspace method that was first used in voice recognition. It eliminates the unwanted information between different pronunciations of the same word due to environmental influences, temporal changes, and personal differences. As a feature extraction method, CVA is based on calculating a unique vector that is invariant to each word called a "Common Vector". It represents common or invariant properties of the word class in the training set. For each word class which is composed from the different pronunciations of the same word caused by the reasons mentioned above, the method is applied separately, taking only into account the within-class distributions. Thus, for each word class, by subtracting a single common feature vector that defines word is obtained. This common vector spans the indifference subspace of the word class [11].

In many pattern classification problems, number (N) of the objects in the training set of each class are equal or less than the size (m) of the attribute space $(N \leq m)$. This condition is called insufficient data case [24]. In this case $\{X_{c1}, X_{c2}, \ldots, X_{cn}\}$ and $X_{ci} = (x_{i1}^c, x_{i2}^c, \ldots, x_{im}^c)$, where c is the class numerator, while n is the total object number in class c. Within-class covariance matrix can be found as

$$C_c = \frac{1}{n} \sum_{i=1}^{n} (X_{ci} - \bar{X}_c)(X_{ci} - \bar{X}_c)^{\mathrm{T}}. \tag{7}$$

Here, $\bar{X}_c = \frac{1}{n}\sum_{i=1}^{n} X_{ci}$ is the mean vector of class c. Then the eigenvalue-eigenvector decomposition is applied to the class covariance matrix obtained in (7). From this decomposition, the m-dimensional feature space spanned by all the eigenvectors can be divided into two subspaces, such as $(n-1)$-dimensional difference subspace and $(m-n+1)$-dimensional orthogonal indifference subspace. In the insufficient data case, where $(n \leq m)$, the $(m-n+1)$ eigenvalues will be zero and the feature subspace spanned by the eigenvectors corresponding to the zero eigenvalues is called indifference subspace and the difference subspace is the compliment of the indifference subspace. Then for the insufficient data case a common vector for each class can be obtained by projecting any vector in that class onto the indifference feature subspace. It is expressed as

$$vec_{com}^c = X_{ci}E_c. \tag{8}$$

The columns of the matrix

$$E_c = \left\{ \begin{array}{ccc} \uparrow & & \uparrow \\ e_1^c & \dots & e_k^c \\ \downarrow & & \downarrow \end{array} \right\}$$

where $k = (m-n+1)$, are the eigenvectors corresponding to zero eigenvalues of the c-th class covariance matrix, or in other words — the vectors that span the indifference feature subspace.

The reasoning mentioned above explains the determination of the common vector which is unique in insufficient data case, where $(n \leq m)$. If the considered case is sufficient data case $(m \leq n)$, in short, if the number n of objects is larger than the size of the attribute space dimension m, the zero eigenvalues will not be obtained from the eigenvalue eigenvector decomposition of the class covariance matrix. This means that the indifference feature subspace has disappeared, so the common vector for the class will just be the zero vector.

However, Gulmezoglu et al. showed that, in the case of sufficient data, indifference feature subspace does not disappear. Because some eigenvalues of the within-class distribution matrix are too small due to the others, and the class common vector still can be obtained [24]. Class common vectors are obtained by projecting the average vector of the class onto the indifference feature subspace that is spanned by the eigenvectors whose value is sufficiently small than the others.

3. Experimental studies

Two different experimental studies are performed in this study for two types of PV fault diagnosis, where one of them is degradation of series resistance and the other one is partial shading. PV faults for the application are simulated using the PSIM program. The data obtained by the simulation in PSIM are analyzed with CVA method proposed in this study and also PCA method is applied to compare the analysis results.

Both of the methods CVA and PCA are supervised learning algorithms as they require a training process. In the training process, the classification model is constructed using the training data set. Once the model is constructed, the performance of the test set is evaluated using the classification model. In the simulated database, the data matrix is composed according to

$$X = [V\,I\,P] \tag{9}$$

The data matrix X contains three attributes, i.e., the voltage V, the current I and the power P produced by the PV system.

3.1. The series resistance degradation

In the first experimental study series resistance degradation fault diagnosis with CVA is carried out. In order to evaluate the performance of CVA method over PCA, the diagnosis is performed with using also PCA. The database is constructed from the simulation results of healthy mode and three degraded resistance mode, i.e.,

$$R_{s_{add}} = 50\%R_s, \quad R_{s_{add}} = R_s, \quad \text{and} \quad R_{s_{add}} = 2R_s.$$

As it can be seen from the behaviour of I–V curves in Fig. 4, there is a redundancy in the data among the healthy condition and the three faulty conditions. Each of the I–V curves are composed of 101 samples that are recorded in one minute. These 101 samples are enough to sweep a complete I–V characteristic of the system. The data matrix is constructed as in the form given in (9). Thus, a data matrix is a 404×3 dimensional matrix, where each row of its corresponds to a sample, and each column corresponds to an attribute such as V, I and P, as shown in Fig. 13. Using the data matrix the training set (67% of all data) and the test set (remain data) that are composed from the samples selected randomly from the data matrix in each time are constructed and the classification-diagnosis results are obtained. This procedure is repeated 10 times and the mean of the results is given as the classification (diagnosis) performance of the model.

As it can be seen in Fig. 15, we cannot classify the R_s degradation faults using PCA, because all the different classes, healthy, $R_{s_{add}} = 50\%Rs$, $R_{s_{add}} = R_s$ and $R_{s_{add}} = 2R_s$, are displayed very close to each other in the PCA feature subspace. That is why the different classes cannot be differentiated. Degraded series resistance fault diagnosis was studied. In the literature using PCA also the same results were obtained. Authors however stated that additional data processing or/and additional information should be included or done in their study in order to separate the healthy case from the degraded series resistance [12].

On the contrary, using CVA method in the same fault condition as it is proposed in this study, we obtained extremely high classification rates without including additional data processing or/and additional information as stated. The reason can easily be seen from the CVA feature space given in Fig. 16. The data scatter lines in CVA
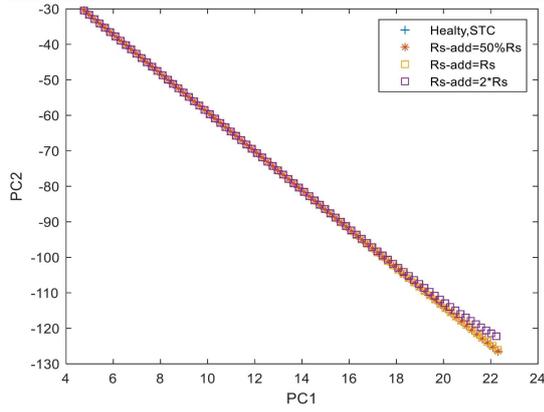
Fig. 15.   Data projected onto the PCA feature space.
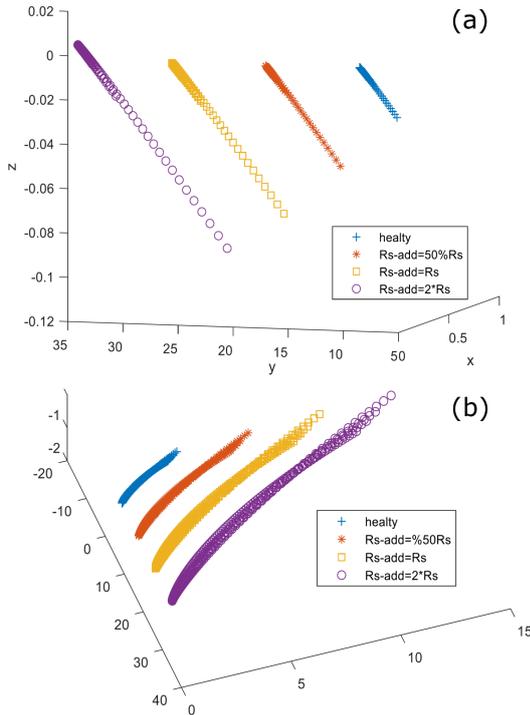
Fig. 16.   (a) Training data, (b) test data projection onto the CVA feature space.

space, corresponding to each condition such as healthy $R_{s_{add}} = 50\%R_s$, $R_{s_{add}} = R_s$ and $R_{s_{add}} = 2R_s$, can easily separable from the others. That is why there are obtained extremely high recognition rates such as 97.5%.

3.2. Partial shading fault mode

In the second experimental study partial shading fault diagnosis with CVA is carried out. In order to evaluate the performance of CVA method over PCA, the diagnosis is performed with using also PCA. The database is constructed using the simulation results of healthy condition and also the simulation results of partial shading conditions given in Figs. 8–11. In the database each partial

TABLE II

Classification percentages in partial shading fault detection.

Method	CVA	PCA
classification performance [%]	99.6	95.4

shading condition (class) corresponds to a class. Each of the class has three relevant I–V curves that are obtained under different irradiation conditions.

As it is in the first experimental study, each of the I–V curves is composed of 101 samples that are recorded in one minute. Thus, each class has a 303×3 dimensional data matrix that includes the $\left[\frac{V}{\eta} \ I \ \frac{P}{\eta}\right]$ measurements taken under different irradiation conditions. Here, η is used for the elimination of the environmental effects and is called efficiency. In total for 5 different working conditions such as one healthy and four partial shading, the total data matrix dimension is 1515×3. Using the data matrix the training set (67% of all data) and the test set (remaining data) composed from the samples selected randomly from the data matrix are constructed and the classification-diagnosis results are obtained. This procedure is repeated 10 times and the mean of the results is given as the classification-diagnosis performance of the model. The classification results are given in Table II.

4. Conclusion

In this paper PV fault detection and classification method, CVA, is proposed for fault diagnosis. In the study two types of fault are taken into account. One of them is series resistance degradation and the other one is partial shading. The CVA method's performance is evaluated with PCA method's performance, which is recently proposed for fault detection.

The data used in the study is simulated using PSIM. For series resistance degradation fault, four simulations that correspond to one healthy and three degraded series resistance, are performed. Using CVA it is shown that in CVA feature space the classes could be differentiated from each other successfully, and 97.5% classification performance is obtained which is extremely high. However, using PCA, series resistance degradation cannot be detected as shown in PCA feature space. In the second fault condition, partial shading, CVA also gives higher classification performances according to PCA. Using CVA method we obtained 99.6% classification rate, while using PCA method the classification rate obtained is 95.4%. These results are not surprising because PCA method eliminates the useful information required to differentiate the classes from each other in transition to feature space but CVA method keeps the specific and distinctive information for each class. Therefore, PCA is successful for dimension reduction rather than classification while CVA is successful in classification.

In addition, the fault detection and classification method proposed in this paper, only requiring the available PV voltage, current, and power values, is not sensitive to the weather conditions and also the sudden variations in the environmental conditions such as temperature and solar irradiation. Moreover, it performs the classification independently of any particular PV size.

References

[1] J.P. Ram, N. Rajasekar, M. Miyatake, *Renew. Sustain. Energy Rev.* **73**, 1138 (2017).

[2] M.R. Maghami, H. Hizam, C. Gomes, M.A. Radzi, M.I. Rezadad, S. Hajighorbani, *Renew. Sustain. Energy Rev.* **59**, 1307 (2016).

[3] A. Mellit, G.M. Tina, S.A. Kalogirou, *Renew. Sustain. Energy Rev.* **91**, 1 (2018).

[4] Y. El Basri, M. Bressan, L. Seguier, H. Alawadhi, C. Alsonso, *Sol. Energy* **116**, 247 (2015).

[5] J.A. Tsanakas, L. Ha, C. Buerhop, *Renew. Sustain. Energy Rev.* **62**, 695 (2016).

[6] M. Bressan, Y. El Basri, A.G. Galeano, C. Alonso, *Renew. Energy* **99**, 1181 (2016).

[7] G.M. Tina, F. Cosentino, C. Ventura, in: *Renewable Energy in the Service of Mankind*, Vol. II, Springer, Cham 2016, p. 505.

[8] W. Chine, A. Mellit, in: *Proc. 5th Int. Conf. on Electrical Engineering (ICEE-B)*, 2017, p. 1.

[9] B. Kiriş, O. Bingöl, R. Şenol, A. Altintaş, *Acta Phys. Pol. A* **130**, 55 (2016).

[10] Z. Erdem, *Acta Phys. Pol. A* **132**, 1128 (2017).

[11] M.B. Gulmezoglu, V. Dzhafarov, M. Keskin, A. Barkana, *IEEE Trans. Speech Audio Process.* **7**, 620 (1999).

[12] S. Fadhel, C. Delpha, D. Diallo, I. Bahri, A. Migan, M. Trabelsi, M.F. Mimouni, *Solar Energy* **179**, 1 (2019).

[13] S. Fadhel, A. Migan, C. Delpha, D. Diallo, I. Bahri, M. Trabelsi, M. Faouzi-Mimouni, in: *Proc. 2018 IEEE Int. Conf. on Industrial Technology (ICIT)* 2018, p. 927.

[14] M.F. Harkat, G. Mourot, J. Ragot, *Process Control* **16**, 625 (2006).

[15] M.H. Gharavian, F. Almas Ganj, A.R. Ohadi, H.H. Bafroui, *Neurocomputing* **121**, 150 (2013).

[16] J. Harmouche, C. Delpha, D. Diallo, *Signal Process.* **94**, 278 (2014).

[17] J. Harmouche, C. Delpha, D. Diallo, *Signal Process.* **109**, 334 (2015).

[18] M.B. Gulmezoglu, V. Dzhafarov, A. Barkana, *IEEE Trans. Speech Audio Process.* **9**, 655 (2001).

[19] D.F. Morrison, Multivariate Statistical Methods, McGraw-Hill, New York 1967, p. 156.

[20] M.G. Villalva, J.R. Gazoli, E.R. Filho, *IEEE Trans. Power Electron.* **24**, 1198 (2009).

[21] H.L. Tsai, C.S. Tu, Y.J. Su, in: *Proc. World Congress on Engineering and Computer Science, San Francisco (CA)*, Eds. S.I. Ao, C. Douglas, W.S. Grundfest, L. Schruben, J. Burgstone, Newswood Ltd, 2008, p. 846.

[22] Powersim Inc., *PSIM Tutorial, "How to Use Solar Module Physical Model"*, 2016.

[23] M. Turk, A. Petland, *J. Cogn. Neurosci.* **3**, 72 (1991).

[24] M.B. Gulmezoglu, V. Dzhafarov, R. Edizkan, A. Barkana, *Computer Speech Lang.* **21**, 266 (2007).