A statistical-based approach for fault detection and diagnosis in a photovoltaic system

Elyes Garoudja, Fouzi Harrou, *Member, IEEE,* Ying Sun, Kamel Kara, Aissa Chouder, Santiago Silvestre

***Abstract*— This paper reports a development of a statistical approach for fault detection and diagnosis in a PV system. Specifically, the overarching goal of this work is to early detect and identify faults on the DC side of a PV system (e.g., short-circuit faults; open-circuit faults; and partial shading faults). Towards this end, we apply exponentially-weighted moving average (EWMA) control chart on the residuals obtained from the one-diode model. Such a choice is motivated by the greater sensitivity of EWMA chart to incipient faults and its low-computational cost making it easy to implement in real time. Practical data from a 3.2 KWp photovoltaic plant located within an Algerian research center is used to validate the proposed approach. Results show clearly the efficiency of the developed method in monitoring PV system status.**

I. INTRODUCTION

Energy based on traditional sources (e.g., fuel oil, gasoline and coal) has some negative impacts on the human general health as well as the ambient environment[1]. Also, to meet the requirements of the COP 21 Paris agreement on climate change, many countries in the world are trying to reduce their dependability on some conventional fossil sources. Towards this end, renewable sources of energy such as solar and wind [2] are very promising alternative resources to replace an important proportion of the traditional forms of energy. Photovoltaic (PV) system, as solar based energy source, is one of the most promising clean energy.

In practice, several faults (e.g., open circuits faults, short circuits faults, hotspot faults, total and partial shading faults) may negatively affect the performance of a PV system and more particularly its DC side, by reducing its energetic yield and its efficiency [3]. Faults in DC side of PV systems are often difficult to avoid and can result in energy loss, system shutdown or even in serious safety concerns [4,5,6].

Elyes Garoudja and Kamel Kara are with the Electronic Department, Blida 1 University Blida, Algeria, (g_lyes@hotmail.fr, k.kara68@gmail.com).
Fouzi Harrou and Ying Sun are with King Abdullah University of Science and Technology (KAUST), Computer, Electrical and Mathematical Sciences and Engineering Division,Thuwal23955900, (fouzi.harrou, ying.sun @kaust.edu.sa).
Aissa Chouder is with the Electrical Engineering Department University of m'sila Ichbilia Street, Algeria (aissachouder@gmail.com).
Santiago Silvestre is with the Electronic Engineering Departmen Univesitat Politécnica de Catalunya Barcelona, Spain (Santiago.silvestre@upc.ed).

Fault detection methods used to detect and diagnose faults in PV systems can be classified into two categories [4]: process-history-based methods and model-based methods. Process-history-based methods are based on implicit empirical models derived from analysis of available data. Such approaches are mainly based on computational intelligence and machine learning methods approaches [7-11]. The methods in this category include artificial neural network [8,9], Bayesian neural network [10] and time domain reflectometry [11]. However, process history based methods require the availability of a relevant dataset which describe both healthy and faulty operating cases of PV system. On the other hand, model-based approaches use analytically computed outputs and compare the values with the measurements and signal an alarm when large differences are detected [12]. Several approaches have been reported in the literature using analytical model based on the one-diode model [12]. Of course, the effectiveness of these model-based fault detection approaches relies on the accuracy of the used models.

In this paper, we present an innovative fault detection approach to monitor the DC side of a PV system. Towards this end, we combine the flexibility of one-diode model with the greater sensitivity of the exponentially-weighted moving average (EWMA) chart to detect small changes [13]. To do so, we generate the residuals, which are the difference between the measured and predicted maximum power point (MPP) current, voltage and power using a one-diode model. Then, EWMA chart is used to monitor residuals to reveal any abnormality. The proposed fault detection strategy has been validated using data from a 9.54 kWp PV plant located in CDER in Algiers, Algeria.

The remainder of this paper is organized as follows. In Section II, the single-diode model is briefly reviewed, and Section III introduces the EWMA chart and its use in fault detection. The proposed scheme is described in Section VI. Simulation results and discussion are given in Section V. Finally, Section VI concludes this paper.

II. SINGLE DIODE MODEL:

The one-diode model is the most common used model to predict energy production in photovoltaic cells. It is based on modelling the solar cell as a light generated current source connected in parallel with a diode, and a series and a parallel

resistance accounting for resistive losses (Figure 1).

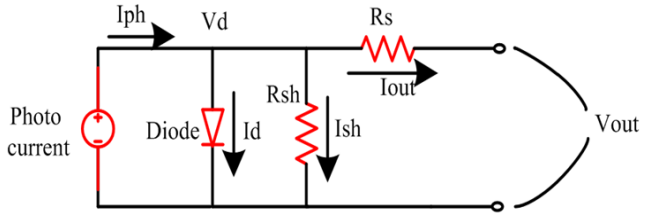

Fig.1. One-diode model of the solar cell/PV module.

The one-diode model is defined to determine the resulting current and voltage, given the parameters of a solar cell.

$$I = I_{ph} - I_0\left(\exp\left(\frac{q(V+R_sI)}{nKT}\right)-1\right)-\frac{V+R_sI}{R_{sh}} \qquad (1)$$

where I and V denote the solar cell's generated current and voltage respectively; I_{ph}: the photo generated currents; I_0: the dark saturation currents; n: the diode's ideality factor; R_s and R_{sh} are the series and shunt resistances respectively; K: Boltzmann constant ($1.38064852 \times 10^{-23}$ J K^{-1}), T: The cell temperature and q is the electronic charge ($1.60217662 \times 10^{-19}$ C).

As it can be seen from (1), the accuracy single diode model is heavily based on five unknowns electrical parameters (I_{ph}, I_0, n, R_s and R_{sh}), hence the great importance of identifying correctly their corresponding values.

III. EWMA Control Chart Theory:

The EWMA is constructed based on exponential weighting of available observations, a design that provides improved sensitivity to small changes in the mean of a multivariate process [14]. The EWMA chart has been originally introduced by Roberts et al in [15], then it has been significantly applied in time series analysis [16-19]. Assuming that $\{x_1, x_2, \ldots, x_n\}$ are the monitored process's collected individual observations, therefore, the statistics of EWMA control chart is computed as follow:

$$\begin{cases} z_t = \lambda x_t + (1-\lambda) z_{t-1} & \text{if } t > 0 \\ z_0 = \mu_0 & \text{if } t = 0 \end{cases} \qquad (2)$$

where μ_0 and λ are the mean of the free-fault process's data and the forgetting parameters respectively; x_t is actual time observation of the monitored process and z_t denotes the EWMA statistic (EWMA's output). The forgetting parameter, $\lambda \in (0, 1]$ determines how fast EWMA forgets the data history.

Indeed, small value of the forgetting parameter λ is suitable to be used to detect relatively small change in the process mean, while a large value of this parameter should be chosen to detect relatively large change in process mean [16,19].

Using this control chart, the process under monitoring will be declared out of control at time *t* if its monitored statistic z_t is outside the control limit, otherwise, it will be considered under control. Where LCL and UCL are EWMA's lower and upper control limits respectively, and they are computed as follow:

$$\begin{cases} LCL = \mu_0 - L\sigma_{z_t} \\ UCL = \mu_0 + L\sigma_{z_t} \end{cases} \qquad (3)$$

where μ_0 is the mean of the free-fault dataset; L is the control limits' width and which defines the limit of confidence (usually L=3; which corresponds to a false alarm rate of 0.27%). σ_{z_t} is the standard deviation and it is computed as:

$$\sigma_{z_t} = \sigma_0\sqrt{\frac{\lambda}{(2-\lambda)}\left(1-(1-\lambda)^{2t}\right)} \qquad (4)$$

σ_0 defines the free-fault dataset's standard deviation.

IV. Fault Detection And diagnosis Strategy:

The procedure of using EWMA for fault detection in PV system is summarized as follows:

(a) Compute residuals using fault-free data:
In this step, we compute residuals, which are defined as difference between the measured and predicted data obtained from simulated model.

$$Res(X) = X_{Real} - X_{pred} \qquad (5)$$

where: Res(X) is the residual of X, X_{Real} and X_{Pred} are the real measurement and the prediction values of X, respectively. The real measured data are collected from the real system under operation and their predicted values are obtained from a simulated model via a PSIM™/Matlab™ Co-simulation. Indeed, to simulate the PV array via the co-simulation, we first identify the five electrical parameters [I_{ph}, I_0, n, R_s, R_{sh}] of a PV module. Then, we construct the simulation model, and compute the residuals of MPP coordinates (I_{mpp}, V_{mpp} and P_{mpp}). This step 1 can be summarized as follow:

- Identify the five unknown electrical parameters of the employed PV module based on an efficient heuristic optimization technic, Artificial Bee Colony (ABC) algorithm [20-22].

- Simulate the employed PV array under PSIM^{TM}/Matlab^{TM} co-simulation for a free-faulty (healthy) operating case and store the MPP coordinates of current I_{mpp}, Voltage V_{mpp} and Power P_{mpp} as the predicted dataset.

- Compute the residuals of the three stored attributes (Res(Impp); Res(Vmpp); Res(Pmpp)) based on the real measured and the predicted data.

(b) Compute EWMA control limits:

In the second step, we compute EWMA control limits for each variable separately (i.e., residuals of DC output current, Res(I_{mpp}), voltage, Res(V_{mpp}), and power, Res(P_{mpp})). The control limits for each are denotes as follows (LCL_{Impp}, $UCL_{Impp,}$ LCL_{Vmpp}, UCL_{Vmpp}, LCL_{Pmpp}, $UCL_{Pmpp,,}$)

(c) Compute EWMA statistic for a new data:

In this phase, testing MPP data of current, voltage and power from the monitored PV system are used to generate residual vectors of current, voltage and power using (5). Then, we compute EWMA monitoring statistics (Z_{Impp}, Z_{Vmpp} and Z_{Pmpp}) based on the previously computed residual vectors using Equation (2).

(d) Check the fault occurrence and diagnose its nature:

When a new sample is available, calculate EWMA based on the ODM model. If EWMA statistics exceeds the confidence limit obtained in the previous step, then a fault is detected. To identify the type of fault we the procedure summarized in Figure 2.

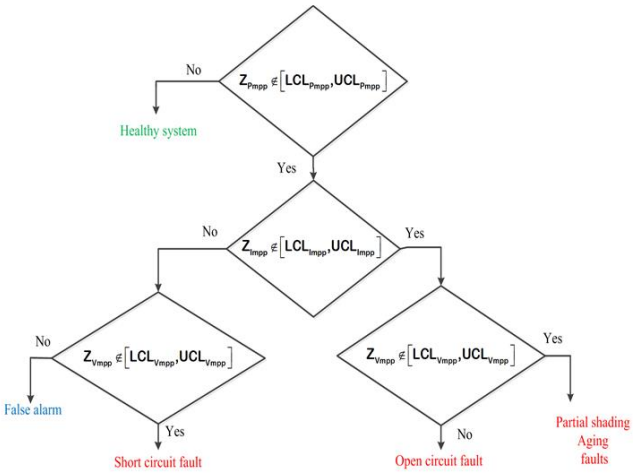

Fig. 2. Fault detection and diagnosis strategy

V. Simulation Results:

The aim of this section is to assess the efficiency of the proposed method. Toward this end, real meteorological conditions of works (Temperatures (T) and Irradiance (G)) and MPP coordinates (I_{mpp}, V_{mpp}, and P_{mpp}) collected from an actual 3.2 KWp PV plant located within an Algerian Research center (Centre du Développement des Energies Renouvelables (CDER) of Algiers) [23] have been used. This PV plant is composed of two parallel strings of fifteen PV modules type Isofoton 106W-12V for each string. The real collected data corresponds to a free-faulty (healthy) system's MPP data of June 24, 2008 day.

As described above, the first step consists in computing the MPP coordinates' residuals (Res (I_{mpp}), Res (V_{mpp}) and Res(P_{mpp})) which is denoted as the subtraction of the real measure value to the predicted based simulation model value of the corresponding attributes.

In this study, the used simulation model is a PSIM^{TM}/Matlab^{TM} Co-simulation model of the 3.2 KWp PV plant of CDER. However, to reach the simulation goal, the five unknown electrical parameters (I_{ph}, I_0, n, R_s and R_{sh}) of its PV module (isofoton 106w-12v) should be accurately identified. To identify unknown electrical parameters of the isofoton 106w-12v PV module, we used a real measured I-V characteristic and an efficient heuristic optimization Algorithm, Artificial Bee Colony (ABC) [20]. The selected parameters are given in table I:

TABLE I. Isofoton 106W-12V PV modules' identified parameters

Parameters	I_{ph}[A]	I_0[A]	n	R_s[Ω]	R_{sh}[Ω]	RMSE
Values	6.54	1.11e-05	1.66	0.147	202.6	0.015

where RMSE is the root mean square error between the real measured current, and the predicted based ABC algorithm.

Once the five parameters have been identified, the employed PV plant's model is validated. Validation step is achieved by introducing the identified parameters to the model, then simulate this last under the same meteorological conditions (T and G) as the measured ones. Finally, the real measured MPP power is compared to the simulated one, as can be appeared in Figure 3.

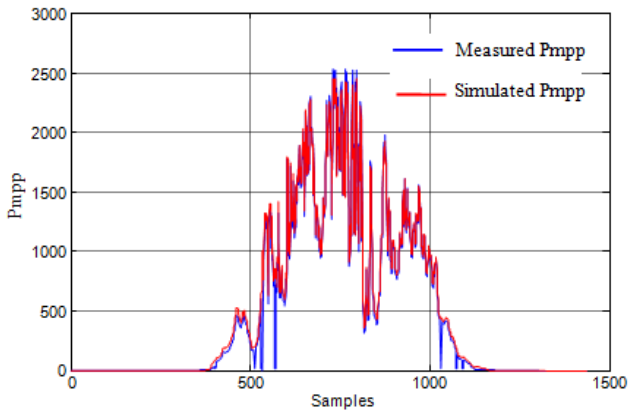

Fig. 3. Real measured versus simulated MPP power.

It can be seen from Figure 3 that the constructed model describes the real behavior of the monitored PV array.

After that, the MPP coordinates' residuals of the free-fault PV system are computed based on (5) using the real measured and the predicted data.

Based on these free-fault system's residuals, EWMA's lower and upper control limits, LCL and UCL, for each one of the three attributes (I_{mpp}, V_{mpp} and P_{mpp}) are computed using (3).

Free-fault system's monitoring statistics should be also calculated based on (2) to check if they are inside the control limits (LCL and UCL).

Simulation results in terms of free-fault system's MPP residuals computations are depicted in *Fig. 4.* While, monitoring statistics and theirs corresponding control limits are depicted in *Fig. 5.*

Based on EWMA control chart in monitoring, we should expect that for the three attributes (I_{mpp}, V_{mpp} and P_{mpp}), their monitoring statistics data will lie within lower and upper EWMA's control limits.

Results from *Fig. 5* show clearly the simulation model's aptitude to describe the free-fault operation case. This conclusion has been directly drawn on the fact that the three monitoring statistics are inside their corresponding EWMA control limits.

Up to now, we have computed the MPP coordinates' residuals and theirs corresponding EWMA's control limits of the free-fault operating PV system.

The goal of the third step is to test the aptitude of our method to detect the DC side faults' occurrence. To this end, three faulty operating cases have been simulated: five modules short-circuited in one string; string completely open-circuited from the array; and finally four modules partially-shaded in the first string operating cases.

The simulated MPP coordinates of the faulty operating cases, will be then used to compute the monitoring statistics for each faulty operating case based on (2).

To test the efficiency of the proposed method, free-fault operation case followed by faulty operation case has been simulated in the following manner:

- Samples from first observation till number 300 correspond to the free-faulty case, while the remaining corresponds to the faulty case (short-circuit and open circuit faults).

- Samples between observation number 150 and 250 corresponds the faulty operations case (partial shading).

Healthy (free-fault) and faulty systems' monitoring statistics are depicted against free-fault's system EWMA control limits for the three faulty operating cases in *Fig. 6*, *Fig. 7* and *Fig. 8* respectively.

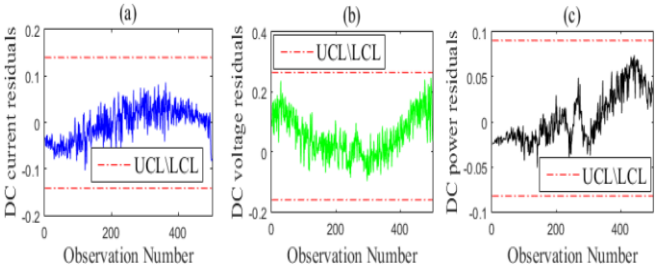

Fig. 4. Free-Fault system's MPP residuals.

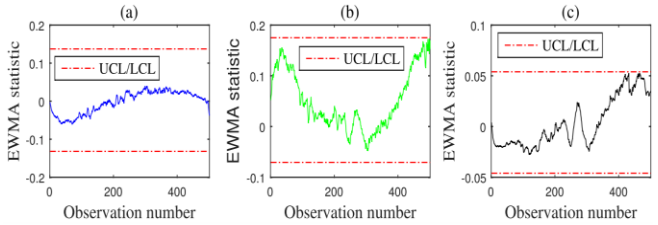

Fig. 5. Monitoring results of EWMA chart for MPP current (a), MPP voltage (b) and MPP power (c) under normal operating conditions.

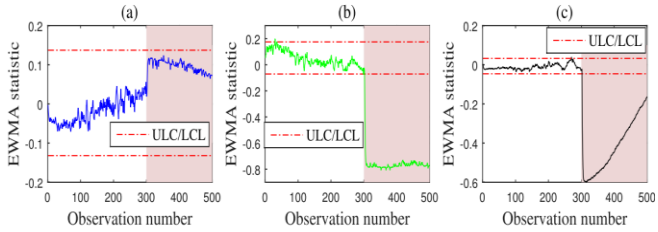

Fig. 6. Monitoring results of EWMA chart for MPP current (a), MPP voltage (b) and MPP power (c) in the presence of five modules short-circuited fault.

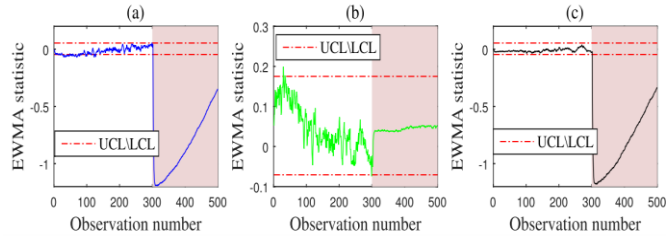

Fig. 7. Monitoring results of EWMA chart for MPP current (a), MPP voltage (b) and MPP power (c) in the presence of string completely open-circuited fault.

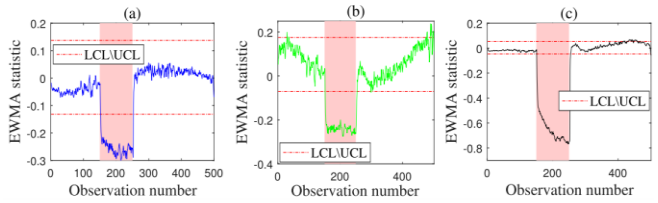

Fig. 8. Monitoring results of EWMA chart for MPP current (a), MPP voltage (b) and MPP power (c) in the presence of four modules partially shaded in the array faulty case.

Finally, faults detections and diagnosis strategy has been proved its high efficiency according to the simulations results.

In fact, since EWMA monitoring statistic of power was inside the free-fault system control limits during its normal operation, then, they move outside these boundaries once a fault has been occurred (*Fig. 6 (c)*, *Fig. 7 (c)*, and *Fig. 8 (c)*).

While, faults diagnosis has been achieved correctly based on the explained topology from flowchart depicted in *Fig. 2* as follows:

Short circuit fault will be declared if the current monitoring statistics (Z_{Impp}) are inside the free-fault system control limits while the voltage ones (Z_{Vmpp}) are outside them (*Fig. 6 (a; b)*)

Open circuits faults will be triggered if the voltage monitoring statistics (Z_{Vmpp}) are inside the free-fault system control limits while the current ones (Z_{Impp}) are outside them *(Fig. 7 (a; b))*

While partial shading of modules faults will be declared when both current and voltage monitoring statistics exceed the free-fault system's control limits (*Fig. 8 (a; b)).*

Finally, the developed method generates a false alarm when Z_{Pmpp} are outside the free-fault system's control limits while the two others (Z_{Impp} and Z_{Vmpp}) are inside their corresponding free-fault system's EWMA Control limits.

VI. CONCLUSION

This paper addressed the problem of fault detection and diagnosis in the DC side of a PV system. An innovative statistical model-based fault detection approach has developed to detect and identify faults in a PV system. First, we used a one diode model (ODM) to generate residuals, which are used as the input for the EWMA monitoring scheme. Indeed, DC power residuals are used for faults detection, while DC current and voltage residuals are used for fault identification. This paper demonstrates through practical data from a grid connected PV system at CDER in Algeria, the effectiveness of the ODM-EWMA to detect and identify faults in a PV system, particularly short-circuit faults, open-circuit faults and partial shading faults.

ACKNOWLEDGEMENT

This publication is based upon work supported by King Abdullah University of Science and Technology (KAUST), Office of Sponsored Research (OSR) under Award No: OSR-2015- CRG4-2582. The authors (Elyes Garoudja and Kamel Kara) thank the SET Laboratory, Department of Electronics, Faculty of Technology, University of Blida 1, Algeria, for continuous support during the study.

REFERENCES

[1] T. B. Johansson, *Renewable energy: sources for fuels and electricity*: Island press, 1993.

[2] N. Panwar, S. Kaushik, and S. Kothari, "Role of renewable energy sources in environmental protection: a review," *Renewable and Sustainable Energy Reviews*, vol. 15, pp. 1513-1524, 2011.

[3] Y. Zhao, R. Ball, J. Mosesian, J.-F. de Palma, and B. Lehman, "Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays," *IEEE Transactions on Power Electronics*, vol. 30, pp. 2848-2858, 2015.

[4] Y. Zhao, "Fault detection, classification and protection in solar photovoltaic arrays," NORTHEASTERN UNIVERSITY, 2015.

[5] B. Brooks, "The Bakersfield Fire: A Lesson in Ground-Fault Protection," *SolarPro Magazine*, pp. 62-70, 2011.

[6] "NC PV DG program SEPA presentation," *DukeEnergy*, pp. 1-14, 2011.

[7] S. Vergura, G. Acciani, V. Amoruso, G. E. Patrono, and F. Vacca, "Descriptive and inferential statistics for supervising and monitoring the operation of PV plants," *IEEE Transactions on Industrial Electronics*, vol. 56, pp. 4456-4464, 2009.

[8] D. D. Nguyen, B. Lehman, and S. Kamarthi, "Performance evaluation of solar photovoltaic arrays including shadow effects using neural network," in *2009 IEEE Energy Conversion Congress and Exposition*, 2009, pp. 3357-3362.

[9] E. Karatepe and T. Hiyama, "Controlling of artificial neural network for fault diagnosis of photovoltaic array," in *Intelligent System Application to Power Systems (ISAP), 2011 16th International Conference on*, 2011, pp. 1-6.

[10] A. M. Pavan, A. Mellit, D. De Pieri, and S. Kalogirou, "A comparison between BNN and regression polynomial methods for the evaluation of the effect of soiling in large scale photovoltaic plants," *Applied energy*, vol. 108, pp. 392-401, 2013.

[11] L. Schirone, F. Califano, U. Moschella, and U. Rocca, "Fault finding in a 1 MW photovoltaic plant by reflectometry," in *Photovoltaic Energy Conversion, 1994., Conference Record of the Twenty Fourth. IEEE Photovoltaic Specialists Conference-1994, 1994 IEEE First World Conference on*, 1994, pp. 846-849.

[12] D. Stellbogen, "Use of PV circuit simulation for fault detection in PV array fields," in *Photovoltaic Specialists Conference, 1993., Conference Record of the Twenty Third IEEE*, 1993, pp. 1302-1307.

[13] J. Lucas and M. Saccucci, "Exponentially weighted moving average control schemes: properties and enhancements," *Technometrics*, vol. 32, no. 1, pp. 1–12, 1990.

[14] N. Zerrouki, F. Harrou, Y. Sun, and A. Houacine, "Accelerometer and camera-based strategy for improved human fall detection," *Journal of medical systems*, vol. 40, no. 12, p. 284, 2016.

[15] Roberts, S.W., "Control chart tests based on geometric moving averages," *Technometrics*.vol.1, no. 3,p. 239–250, 1959.

[16] Kadri, F., Harrou, F., Chaabane, S., Sun, Y., and Tahon, C., "Seasonal ARMA-based SPC charts for anomaly detection: application to emergency department systems,". *Neurocomputing*, no. 173, p. 2102–2114, 2016.

[17] Harrou. F, Nounou. M, Nounou. H, and Madakyaru. M, "PLS-based EWMA fault detection strategy for process monitoring," *J. Loss Prev. Process Ind.* 36:108–119, 2015.

[18] F. Harrou and M. Nounou, "Monitoring linear antenna arrays using an exponentially weighted moving average-based fault detection scheme," Systems Science & Control Engineering: An Open Access Journal, vol. 2, no. 1, pp. 433–443, 2014.

[19] D. C. Montgomery, "Introduction to statistical quality control," John Wiley& Sons, New York, 2005.

[20] D. Karaboga and B. Basturk, "On the performance of artificial bee colony (ABC) algorithm," *Applied soft computing*, vol. 8, pp. 687-697, 2008.

[21] E. Garoudja, K. Kara, A. Chouder, and S. Silvestre, "Parameters extraction of photovoltaic module for long-term prediction using artifical bee colony optimization," in 3rd International Conference on Control, Engineering & Information Technology (CEIT). IEEE, 2015, pp. 1–6.

[22] S. Abou, A. Chouder, K. Kara, and S. Silvestre, "Artificial bee colony based algorithm for maximum power point tracking (MPPT) for PV systems operating under partial shaded conditions," Applied Soft Computing, vol. 32, pp. 38–48, 2015.

[23] A. Chouder and S. Silvestre, "Automatic supervision and fault detection of PV systems based on power losses analysis," Energy Conversion and Management, vol. 51, no. 10, pp. 1929–1937, 2010.