A Regression Algorithm for the Smart Prognosis of a Reversed Polarity Fault in a Photovoltaic Generator

Wail Rezgui, Nadia Kinza Mouss, Leïla-Hayet Mouss, Mohamed Djamel Mouss, Mohamed Benbouzid

A Regression Algorithm for the Smart Prognosis of a Reversed Polarity Fault in a Photovoltaic Generator

Wail Rezgui, Nadia Kinza Mouss, Leïla-Hayet Mouss and Mohamed Djamel Mouss
University of Batna, LAP-lab
Batna, Algeria
r-wail@hotmail.fr

Mohamed Benbouzid
University of Brest, EA 4325 LBMS
Brest, France
Mohamed.Benbouzid@univ-brest.fr

***Abstract*—This paper deals with a smart algorithm allowing reversed polarity fault diagnosis and prognosis in PV generators. The proposed prognosis (prediction) approach is based on the hybridization of a support vector regression (SVR) technique optimized by a k-NN regression tool (K-NNR) for undetermined outputs.**

To test the proposed algorithm performance, a PV generator database containing sample data is used for simulation purposes.

***Keywords*—Photovoltaic generator, SVR, k-NNR, reversed polarity fault, diagnosis, prognosis.**

NOMENCLATURE

PV	=	Photovoltaic;
SVM	=	Support Vector Machines;
SVR	=	Support Vector Regression;
k-NN	=	k-Nearest Neighbor;
k-NNR	=	k-Nearest Neighbor Regression;
X	=	SVM input vector,
Y	=	SVM output vector,
f	=	Linear function;
Φ	=	Kernel function;
w	=	Weight vector;
x	=	Problem variable;
x^*	=	New problem variable;
α	=	Lagrange multipliers.

I. INTRODUCTION

The performance of a photovoltaic generator is unfortunately degraded by the presence of defects which may cause significant power losses; especially reversed polarity fault of some generator components [1-2]. PV generator faults prognosis [3-7] and diagnosis [8-11] can stabilize its performance, reduce its maintenance costs, and ensure its availability and reliability.

In this context, the paper objective is the development of an algorithm for the prognosis of a photovoltaic generator state under reversed polarity faults. Indeed, the paper main contribution is the development of a smart prognosis algorithm Indeed, the paper contributions are twofold: *1)* Development of a fault detection and localization algorithm; in particular reversed polarity at different levels: in a cell, bypass (group level), and blocking diodes (string level). *2)* Development of a smart prognosis algorithm for the characterization of reversed polarity faults, regardless of their localization. The proposed smart algorithm is based on the support vector regression (SVR) [12-15] and k-NN regression (k-NNR) tools [16-25].

II. CLASSICAL DIAGNOSIS ALGORITHM

First of all, this proposes a diagnosis algorithm that is designed for the detection and the localization of reversed polarity faults in PV generator at different levels: a PV cell, bypass, and blocking diodes.

The following characterizes the fault impacts at different level of the PV generator.

A. Faults at the Generator Level

The generator voltage V is good indicator for the fault detection. Indeed, a PV generator is subjected to a reversed polarity in the following case.

$$\begin{cases} V_{PV} < V_{PV\_Healthy} \\ I_{PV} = I_{PV\_Opposite} \\ PHI = 0 \end{cases} \tag{1}$$

Where I is the current and PHI the photo current. The following situation:

$$\begin{cases} -V_{PV\_Healthy} \leq V_{PV} \leq V_{PV\_Healthy} \\ 0 < I_{PV} < I_{PV\_Healthy} \\ PHI \neq 0 \end{cases} \tag{2}$$

means that the generator contains at least one string with a component (cell or bypass diode) subjected to a reversed polarity fault.

B. Faults at a String Level

At the string level, the two following situations can be encountered.

1) Blocking diode under reversed polarity. A blocking diode is subjected to a reversed polarity if

$$\begin{cases} I_{String} = 0 \\ I_{String\_opposite} = 0 \\ PHI \neq 0 \\ V_{String} = V_{String\_open\ circuit} \end{cases} \tag{3}$$

$$\text{or } \begin{cases} I_{String\_Cells} = 0 \\ I_{String\_Cells\_opposite} = 0 \\ PHI = 0 \\ V_{String\_Cells} = 0 \end{cases} \tag{4}$$

2) String module under reversed polarity. A string will contain at least one module under reversed polarity if

$$\begin{cases} 0 < I_{String} < I_{String\_Healthy} \\ -V_{String\_Healthy} \leq V_{String} \leq V_{String\_Healthy} \end{cases} \tag{5}$$

C. Faults at a Module Level

A module is subjected to a reversed polarity if

$$\begin{cases} 0 < I_{Module} < I_{Module\_Healthy} \\ V_{Module} < V_{Module\_Healthy} \end{cases} \tag{6}$$

D. Faults at a Group Level

At the group level, the three following situations can be encountered.

1) A group will contain at least a cell under reversed polarity if

$$\begin{cases} I_{Group} = I_{Group\_Healthy} \\ V_{Group} > 0 \end{cases} \tag{7}$$

In this case, the number of healthy cells is greater than the defective ones.

2) A group will contain at least half of its cells under reversed polarity if

$$\begin{cases} I_{Group} = I_{Cells} + I_{Bypass\ diode} \\ V_{Group} \leq 0 \end{cases} \tag{8}$$

3) A group will be grouped by a bypass diode in reversed polarity if

$$\begin{cases} I_{Group} = I_{Cells} - I_{Bypass\ diode} \\ V_{Group} = 0 \end{cases} \tag{9}$$

III. Intelligent Prognosis Algorithm

A. SVR Algorithm

Support vector machines are supervised learning techniques, for solving classification problems in many fields such as fault diagnosis [26]. In addition SVM have enormous capacities of application in case of regression (SVR).

SVR, as SVM, has also linear and nonlinear cases. For the nonlinear case, it requires a kernel function to transform the data in a new greater dimension space; allowing therefore the data processing in a linear space.

SVR general pattern of regression is as follows

$$\begin{cases} \text{If the problem is linearly separable} \\ \quad f(x) = \langle wx \rangle + b \\ \text{Else} \\ \quad f(x) = \langle w\phi(x) \rangle + b \end{cases} \tag{10}$$

where w and b are the parameters that should maximize the flat maximum of the linear function f (in the nonlinear case after space transformation).

The above problem can therefore be written in terms of linear programming and solved via Lagrange minimization [27-30]. The problem becomes then a regression as follows.

$$\begin{cases} \text{If the problem is linearly separable} \\ \quad f(x) = \left[\langle xx_1 \rangle \langle xx_2 \rangle \ldots \langle xx_N \rangle\right]\alpha + b \\ \text{Else} \\ \quad f(x) = \left[\langle \phi(x)\phi(x_1) \rangle \ldots \langle \phi(x)\phi(x_N) \rangle\right]\alpha + b \end{cases} \tag{11}$$

SVR, like other regression tool, has several disadvantages depending on its use, such as the kernel function choice, its parameters, its algorithmic complexity, etc. This paper is particularly proposing a solution to the problem of value indetermination of some SVR output (that would be predicted). Indeed, it is proposed the use of the k-Nearest Neighbor Regression (k-NNR) to predict an approximate value of an undetermined SVR output.

B. k-NN Regression Algorithm

The k-NNR has been chosen as it provides a set of performance, among them: *1)* No required learning (no required model); *2)* Forecasts based on only all the nearest observations to predict new observation approximate numerical values.

The proposed k-NNR algorithm consists in three main steps.

1) ***Step 1***. Identify the new observation (x^*) collected from the system (the PV generator in our case).

2) ***Step 2***. Calculate the distance between the new observation and all the database previous ones of the database. The Manhattan distance is adopted in our case (Step 3).

3) ***Step 3***. Choose from the previously calculated distances the minimum one and select it index:

$$MIN = \text{index}\left(\min\begin{bmatrix}\left|x^{*}-x_{1}\right| \\ \left|x^{*}-x_{2}\right| \\ \vdots \\ \left|x^{*}-x_{n}\right|\end{bmatrix}\right) = \begin{bmatrix}\text{index}_{1} \\ \text{index}_{2} \\ \vdots \\ \text{index}_{m}\end{bmatrix}, \text{with } m \le n \quad (12)$$

If there is a single minimum, there is only one approximated output. The output of this new observation is equal to the numerical value of the nearest one. Otherwise, if there are several minimum distances, this means that there is more than one approximated output. In this case, the output of the new observation is the mean value of the nearest outputs.

If the achieved distance is very large (defined by the expert), it belongs then to the distance discharges class. This is a limitation of the proposed algorithm.

C. The Proposed Smart Algorithm

In this paper, it is proposed a smart algorithm allowing smart prognosis of a PV generator condition. In particular, it is proposed to optimize the accuracy of its output prediction according to given inputs.

This algorithm is considered smart as it results from the hybridization of the SVR technique (artificial intelligence family) optimized by the k-NNR tool that is aimed to solve the SVR problem of undetermined outputs. The hybridization of these two methods leads to the following formulation.

$$\begin{cases} \text{If the problem is linearly separable} \\ f(x) = \left[(1-\theta)\left(\left[\langle xx_{1}\rangle\langle xx_{2}\rangle\ldots\langle xx_{N}\rangle\right]\alpha + b\right) + \theta\left(\dfrac{\sum_{p=1}^{m}\left(\left[\left\langle x_{\left[\text{index}_{p}\right]}x_{1}\right\rangle\left\langle x_{\left[\text{index}_{p}\right]}x_{2}\right\rangle\ldots\left\langle x_{\left[\text{index}_{p}\right]}x_{N}\right\rangle\right]\alpha + b\right)}{m}\right)\right] \\ \text{Else} \\ f(x) = \left[\begin{array}{c}(1-\theta)\left(\left[\langle\phi(x)\phi(x_{1})\rangle\langle\phi(x)\phi(x_{2})\rangle\ldots\langle\phi(x)\phi(x_{N})\rangle\right]\alpha + b\right) \\ +\theta\left(\dfrac{\sum_{p=1}^{m}\left(\left[\left\langle\phi\left(x_{\left[\text{index}_{p}\right]}\right)\phi(x_{1})\right\rangle\left\langle\phi\left(x_{\left[\text{index}_{p}\right]}\right)\phi(x_{2})\right\rangle\ldots\left\langle\phi\left(x_{\left[\text{index}_{p}\right]}\right)\phi(x_{N})\right\rangle\right]\alpha + b\right)}{m}\right)\end{array}\right]\end{cases} \quad (12)$$

$\theta = 0$ for determined SVR results, else $\theta = 1$.

IV. Simulations Results

A. Faulted PV Generator Characterization

To illustrate the performance of the fault detection and diagnosis algorithms several simulations have been carried-out on a typical PV generator.

The main simulation results are shown by Fig. 1 to 3. Figure 1 clearly illustrates the PV generator power performance with cells under reversed polarity. In particular, this figure shows that the power decreases proportionally to the number of defective cells. In this context, it should be noted that the current is independent of the number of defective cells but the voltage will be null in the case where half of the cells are defective. Moreover, if the number of defective cells is greater that the healthy ones, the PV generator absorbs power.

Figure 2 shows the influence of bypass diodes under reversed polarity on the generated power. These defective diodes have a greater impact (compared to defective cells) on the PV generator performance. Indeed, defective bypass diodes affect a group voltage.

Finally, Fig. 3 illustrates the influence of blocking diodes under reversed polarity on the PV generator operation. It is clearly shown that when this type of fault occurs, the generated power degradation is faster compared to the case with defective bypass diodes. Indeed, a defective blocking diode will affect a full string (affect the current flow and behaves like an open-circuit).

B. Smart Algorithm Tests

The proposed smart prognosis algorithm is tested using 50 samples related to a typical PV generator.

Figure 4 illustrates the prediction performance in terms of absolute and relative errors. It is clearly shown that the proposed hybrid SVR + KNNR method achieves the best prediction performance.

Figure 5 illustrates the second criterion used for the prediction performance evaluation: variance and standard deviation.

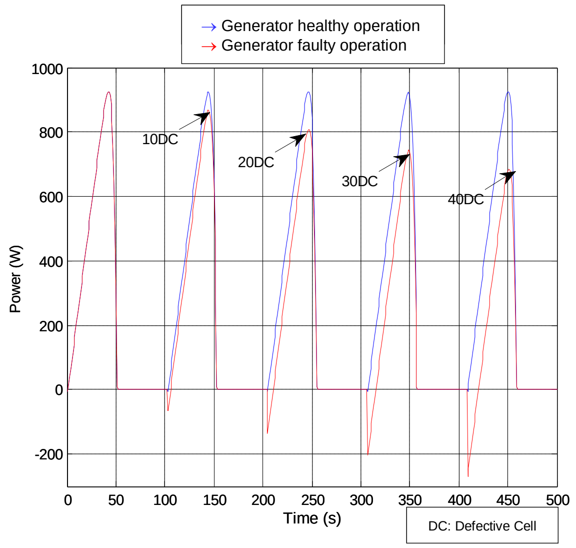

Fig. 1. Cells reversed polarity influence on the PV generator operation.

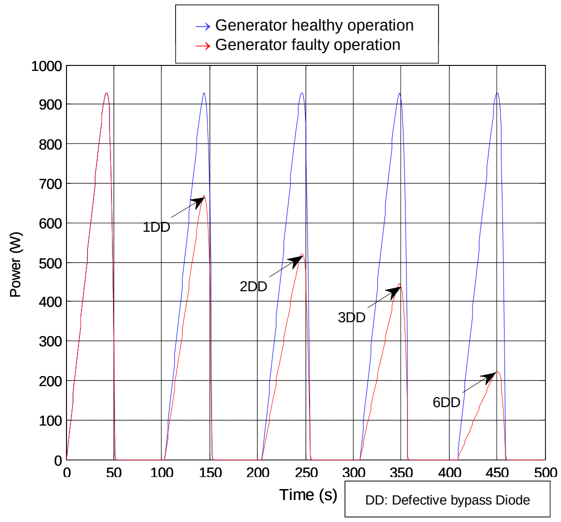

Fig. 2. Bypass diodes reversed polarity influence on the PV generator operation.

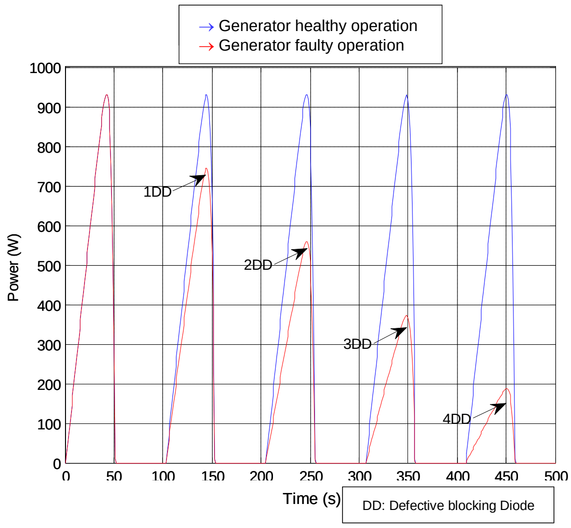

Fig. 3. Blocking diodes reversed polarity influence on the PV generator operation.

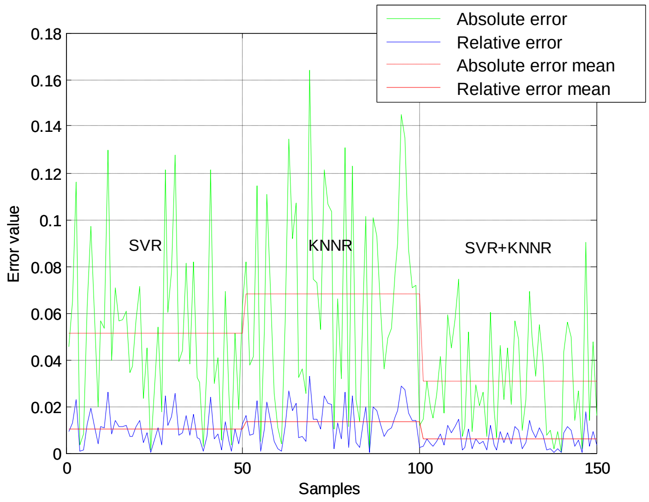

Fig. 4. Absolute and relative errors *vs* prediction tools.

The achieved results clearly show that the proposed hybrid SVR + KNNR method leads to the lowest variance and standard deviation: *this means that the prediction results of the proposed method are more homogeneous*.

The third evaluation criterion used is the confidence interval. In this case, Fig. 6 illustrates the prediction performance and clearly shows that the proposed hybrid SVR + KNNR method has the most optimized confidence interval allowing therefore more reliable prediction.

V. Conclusion

This paper dealt with a smart algorithm allowing reversed polarity fault diagnosis and prognosis in PV generators. The proposed prognosis (prediction) approach is based on the hybridization of a support vector regression (SVR) technique optimized by a k-NN regression tool (K-NNR) for undetermined outputs.

The proposed smart prognosis algorithm was tested using 50 samples related to a typical PV generator.

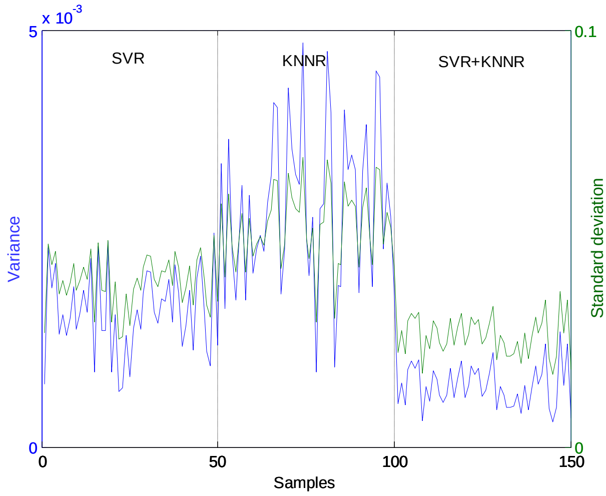

Fig. 5. Variance and standard deviation *vs* prediction tools.

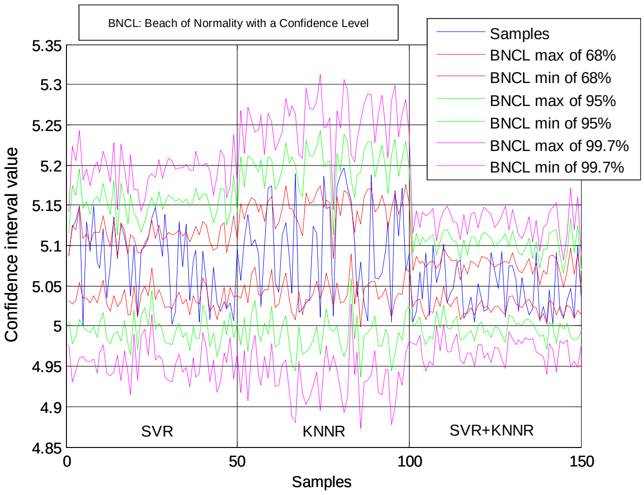

Fig. 6. Confidence interval *vs* prediction tools.

The analysis of the simulation results shows that the proposed hybrid prognosis technique achieves the best prediction performance with more homogeneous and reliable prediction results.

References

[1] V.S. Hosseini, S.Y. Woo, S. Ghanad-Tavakoli, R.N Kleiman, J.S. Preston and G.A. Botton, "Atomic-resolution study of polarity reversal in GaSb grown on Si by scanning transmission electron microscopy," *Journal of Applied Physics*, vol. 112, n°9, pp. 093101-1-093101-9, November 2012.

[2] G. Liu, L. Li, F. Ji, W. Li, Y. Sun, B. Li and J. Li, "Analysis of transient electric field and charge density of converter transformer under polarity reversal voltage," *IEEE Trans. Magnetics*, vol. 48, n°2, pp. 275-278, February 2012.

[3] B.E. Olivares, M.A. Cerda Munoz, M.E. Orchard and J.F. Silva, "Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena," *IEEE Trans. Instrumentation and Measuremen*t, vol. 62, n°2, pp. 364-376, February 2013.

[4] Y. Wang, M. Simon, P. Bonde, B.U. Harris, J.J. Teuteberg, R.L. Kormos and J.F. Antaki, "Prognosis of right ventricular failure in patients with left ventricular assist device based on decision tree with SMOTE, "*IEEE Trans. Information Technology in Biomedicine*, vol. 16, n°3, pp. 383-390, May 2012.

[5] J. Son, Q. Zhou, S. Zhou, X. Mao and M. Salman, "Evaluation and comparison of mixed effects model based prognosis for hard failure," *IEEE Trans. Reliability*, vol. 62, n°2, pp. 379-394, June 2013.

[6] D.M. Sow, J. Sun, A. Biem, J. Hu, M.L. Blount and S. Ebadollahi, "Real-time analysis for short-term prognosis in intensive care," *IBM Journal of Research and Development*, vol. 56, n°5, pp. 3:1-3:10, September-October 2012.

[7] P. Arpaia, C. Manna, G. Montenero and G. D'Addio, "In-time prognosis based on swarm intelligence for home-care monitoring: A case study on pulmonary disease," *IEEE Sensors Journal*, vol. 12, n°3, pp. 692-698, March 2012.

[8] Y. Hirata, S. Noro, T. Aoki and S. Miyazawa, "Diagnosis photovoltaic failure by simple function method to acquire I–V curve of photovoltaic modules string," *in Proceedings of the 2012 IEEE PVSC*, Austin (USA), pp. 1340-1343, June 2012.

[9] M. Gonzalez, B. Raison, S. Bacha and L. Bun, "Fault diagnosis in a grid-connected photovoltaic system by applying a signal approach," *in Proceedings of the 2011 IEEE IECON*, Melbourne (Australia), pp. 1354-1359, November 2011.

[10] T. Jianeng, Z. Yongqiang and W. Wenshan, "Fault diagnosis method and simulation analysis for photovoltaic array," *in Proceeding of the 2011 IEEE ICECE*, Yichang (China), pp. 1569-1573, September2011.

[11] X. Xu, H. Wang and Y. Zuo, "Method for diagnosing photovoltaic array fault in solar photovoltaic system," *in Proceeding of the 2011 IEEE APPEEC*, Wuhan (China), pp. 1-5. March 2011.

[12] J. Nong, "Parameters selection and noise estimation of SVM regression," *in Proceedings of the 2012 IEEE CSO*, Harbin (China), pp. 118-122, June 2012.

[13] J. Li, S. West and G. Platt, "Power decomposition based on SVM regression," *in Proceedings of the 2012 IEEE ICMIC*, Wuhan (China), pp. 1195-1199, June 2012.

[14] J. Zhang, K. Li, G.W. Irwin and W. Zhao, "A regression approach to LS-SVM and sparse realization based on fast subset selection," *in Proceedings of the 2012 IEEE WCICA*, Beijing (China), pp. 612-617, July 2012.

[15] J. Zabalza, J. Ren, C. Clement, G. Di Caterin and J. Soraghan, "Embedded SVM on TMS320C6713 for signal prediction in classification and regression applications," *in Proceedings of 2012 IEEE EDERC*, Amsterdam (Netherlands), pp. 90-94, September 2012.

[16] F. Khelifi and J. Jiang, "K-NN regression to improve statistical feature extraction for texture retrieval," *IEEE Trans. Image Processing*, vol. 20, n°1, pp. 293-298, January 2011.

[17] A.J. Eronen and A.P Klapuri, "Music tempo estimation with K-NN regression," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 18, n°1, pp. 50-57, January 2010.

[18] H. Hirose, Y. Soejima and K. Hirose, "NNRMLR: A combined method of nearest neighbor regression and multiple linear regression," *in Proceedings of the 2012 IEEE IIAIAAI*, Fukuoka (Japan), pp. 351-356, September 2012.

[19] O. Kramer, "Dimensionality reduction by unsupervised k-nearest neighbor regression," *in Proceedings of the 2012 IEEE ICMLA*, Honolulu (USA), pp. 275-278, December 2011.

[20] G. Biau, L. Devroye, V. Dujmovic and A. Krzyzak, "An affine invariant k-nearest neighbor regression estimate," *in Proceedings of the 2012 IEEE ISIT*, Cambridge (USA), pp. 1445-1447, July 2012.

[21] S. Li, Z. Shen and G. Xiong, "A k-nearest neighbor locally weighted regression method for short-term traffic flow forecasting," *in Proceedings of the 2012 IEEE ITSC*, Anchorage (USA), pp. 1596-1601, September 2012.

[22] J. Wu, "A novel artificial neural network ensemble model based on k-nearest neighbor nonparametric estimation of regression function and its application for rainfall forecasting," *in Proceedings of the 2009 IEEE CSO*, Sanya (China), pp. 44-48, April 2009.

[23] M. Luo, Y. Zheng and S. Liu, "Data-based fault-tolerant control of the semiconductor manufacturing process based on k-nearest neighbor nonparametric regression," *in Proceedings of the 2012 IEEE WCICA*, Beijing (China), pp. 3008-3012, July 2012.

[24] Z.W. Yuan and Y.H. Wang, "Research on k nearest neighbor non-parametric regression algorithm based on KD-tree and clustering analysis," *in Proceeding of the 2012 IEEE ICCIS*, Chongqing (China), pp. 298-301, August 2012.

[25] G. Morales-España, J. Mora-Flórez and G. Carrillo-Caicedo, "A complete fault location formulation for distribution systems using the k-Nearest Neighbors for regression and classification," *in Proceedings of the 2010 IEEE TDC-LA*, Sao Paulo (Brazil), pp. 810-815, November 2010.

[26] W. Rezgui, L.H. Mouss, N.K. Mouss, M.D. Mouss and M.E.H. Benbouzid, "A smart algorithm for the diagnosis of short-circuit faults in a photovoltaic generator," *in Proceedings of the 2014 IEEE ICGE*, Sfax (Tunisia), pp. 1-5, March 2014.

[27] Z. Lu, J. Sun and K. Butts, "Linear programming SVM-ARMA 2K with application in engine system identification," *IEEE Trans. Automation Science and Engineering*, vol. 8, n°4, pp. 846-854, October 2011.

[28] X. Chen, J. Yang, J. Liang and Q. Ye, "Robust and sparse twin support vector regression via linear programming," *in Proceedings of the 2010 IEEE CCPR*, Chongqing (China), pp. 1-6, October 2010.

[29] Z. Lu, J. Sun and K. Butts, "Application of linear programming SVM-ARMA2K for dynamic engine modeling," *in Proceedings of the 2010 IEEE ACC*, Baltimore (USA), pp. 1465-1470, , June-July 2010.

[30] H. Xie, J.P. Wei and H.L. Liu, "Linear programming regressive support vector machine," *in Proceedings of the 2006 IEEE ICMLC*, Dalian (China), pp. 2196-2199, August 2006.