Neural Principal Component Analysis for ECG Signal Monitoring

Hanen Chaouch Jebril, Khaled Ouni
Engineering National school of Monastir
Automatic and Signal & Image Processing
Monastir, Tunisia
hanenchaouch8@yahoo.com, {khaled_ouni,lotfinabli}@yahoo.fr

ABSTRACT: *In this paper, we address the problem of monitoring the cardiovascular system through the integration of automatic tools. This monitoring is to detect heart diseases starting from the electrocardiographic signal (ECG). In particular, we propose a method for detection of defects in the ECG signal analysis exploring the non linear principal components (NLPCA). The data matrix consists of 528 measurements and 9 variables, these variables are determined from waves and segments of the ECG. The proposed approach allows to reduce the size of the data matrix and find the linear and non-linear relationships between variables. Defect detection is established by the statistical SPE (square predictive error) which is based on residues. Defective variables are found by calculating their contributions, its have the highest contributions.*

Keywords: ECG, Non Linear Principal Component Analysis, Deectt Detection, NLPCA, Auto-associative Network, Algorithms of Learning

Received: 27 April 2013, Revised 1June 2013, Accepted 25 June 2013

1. Introduction

The cardiovascular system is one of the biological processes which took a lot of interest by the previous and current researches. Its analysis and its supervision was establishes through the signal supplied by the electrocardiogram called ECG. Besides analyses workbenches by the expert (cardiologists), the supervision of the cardiovascular system by the introduction of the tools of signal processing of and modeling can give an idea onto the state of the heart (normal or sick). This supervision is establishes by integrating the characteristics of the ECG, its waves (*P*, *QRS* and *T*) and or by its segments. It consists in the screening and the classification of the cardiac arrhythmias. Several works introduced the artificial intelligence for the analysis of the state of the cardiovascular system of others used the combination of various methods (neuron network, fuzzy logic) [1, 2, 3 and 4]. These applied approaches are methods of processing of quantitative and qualitative data.

These computing tools assert themselves at the moment as a new major medical technology, the purpose of which is not to replace the work of the expert but to facilitate it. The supervision of the processes with multi-variable statistical tools is developed during these last years. Among these statistical tools, the principal components analysis is widely used for the surveillance of the industrial and biological processes [5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15]. The PCA is a technique of projection

of the data towards a space of reduced dimension. This method in summer often used given its capacity to capture the linear relations between the variables of the system at the still state. However, this approach has limits to treat data in the form of measures on industrial or biological processes presenting generally non linear characteristics. Often by using of the linear PCA, we cannot indicate the non linear part of variables involved in the data. For that purpose, we need methods with non linear PCA to exploit this non linearity between the parameters for the surveillance of the system [16, 17, 18 and 19].

In this work, we propose an approach of detection of the defects by using a neural principal components analysis. This approach is applied to a signal ECG to set up the pathological state of the cardiovascular system.

2. The Electronicardiogram (ECG)

A cardiac fiber in the course of depolarization can be likened to an electric dipole. At the given moment, the front of the wave of activation trained by all the elementary dipoles creates an electric field which is a function of depolarizes moments. The recording of the temporal evolution of the resultant electric field, made by means of cutaneous electrodes, is called the electrocardiogram of surface.

2.1 Configuration of normal ECG

For every cardiac beating, the electrocardiogram registers successively 5 waves: *P*, *Q*, *R*, and *T* (figure 1). Generally the waves *Q*, *R* and *S* are grouped and we speak rather about the complex QRS [20].

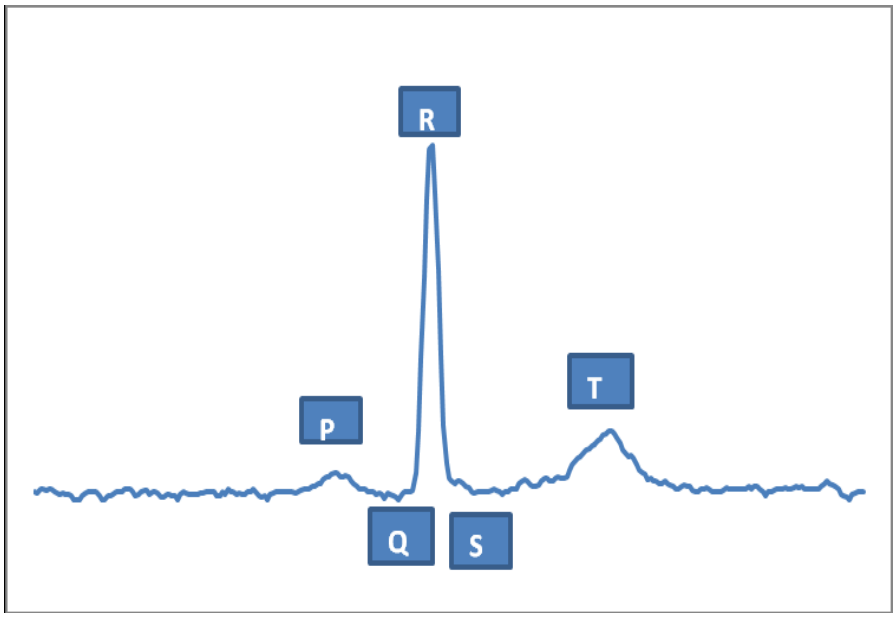

Figure 1. Normal ECG

• **The wave *P*:** It represents the auricular depolarization. its duration is of the order of 90 ms. It is a positive wave the amplitude of which is normally lower or equal in 0.2 mV.

• **The complex QRS:** It corresponds to the ventricular depolarization preceding the mechanical effect of contraction. its normal duration is included between 85 and 95 ms.

• **The wave *T*:** It corresponds to the repolarization of ventricles. It is asymmetric, of a slightly oblique ascending branch and an steeper downward branch. Its amplitude is lower than 2 mm. The normal wave *T* have an amplitude lower than the complex QRS.

2.2 The ECG intervals

The intervals and the segments of the electrocardiographic plan of the important parameters, allow to estimate the normality or not of the space between two electric events.

• **The PR (or P-Q) interval:** Corresponds in time between auricular and ventricular depolarization. It is the weather of propagation

of the wave of depolarization through auricles, the atrio-ventriculaire knot, the beam of HIS and the network of PURKINJE, up to ventricular myocardiques.

• **The *QT* interval:** C orresponds in the time of ventricular systole which goes of the beginning of the ventricles excitement of till the end of its relaxation.

• **The segment *ST*:** It corresponds to ventricular repolarization, phase in the course of which corresponds to the phase of ventriculars cells are quite depolarized: there is no electric extension, the segment is then iso-electric.

• **The RR interval:** Separates the summits of two successive waves R and represents the immediate frequency.

3. Non Linear Prinicpal Components Analysis (NLPCA)

3.1 Principle of NLPCA

The linear principal components analysis is a projection based on statistical tools traditionally used for the reduction of dimension. We consider the matrix of data $Z = [z_1, z_2,..., z_m] \in \Re^{N \times m}$, the non linear principal components analysis z is presented as follow:

$$Z = TP^T + E = \sum_{i=1}^{\ell} t_i p_i + E \qquad (1)$$

with $T = [t_1, t_2,..., t_m]$ is the principal components matrix, $P = [p_1, p_2,..., p_m]$is eigenvectors matrix and E is the matrix of residue.

The PCA identifies linear correlations between the variables of process. On the other hand, the non-linear PCA can extract correlations as linear as non-linear. The generalization is realized by throwing the variables of process on curves or surfaces and not on lines or plans using the same objective, by minimizing the cost function following one:

$$E = \{\| z - \hat{z} \|^2\} \qquad (2)$$

The matrix of data Z can be in term of non linear principal components ℓ where $\ell \prec m$, by the following equation:

$$Z = F(T) + E \qquad (3)$$

where $T = [t_1, t_2,..., t_m]$ is the principal components matrix, F is a non linear function is which corresponds to the matrix of the eigen vectors in case of linear PCA and E is the matrix of residues.

3.2 NLPCA based on Neural Network

The structure of neural network in the figure 2 is applied, by using five series-parallel layers. A first layer of entrance, a layer in the middle called sleeps of constriction which determines the number of principal components and a layer of exit. In a first layer hidden from coding, we resort to a function of non linear transfer sigmoïdale and in the third layer of decoding or reconstruction, we have used a function of non linear transfer sigmoïdale defined by the following equation :

$$\sigma(x) = \frac{1}{1 + e^{-\beta x}} \qquad (4)$$

The number of principal components t_ℓ in the layer of constriction is determinate using the Webb criterion [21] defined as follow:

$$\varepsilon(\ell) = \sqrt{\frac{\| \hat{X} - X \|^2}{\| X - \bar{X} \|^2}} \qquad (5)$$

where $\bar{X}$ A matrix is of which the vectors consist of average vectors of the matrix X. The ℓ number will be increased and will be watched at the same time.

3.3 Defect detection

For the detection of defects, we use the quadratic error called SPE expressed by the following formula:

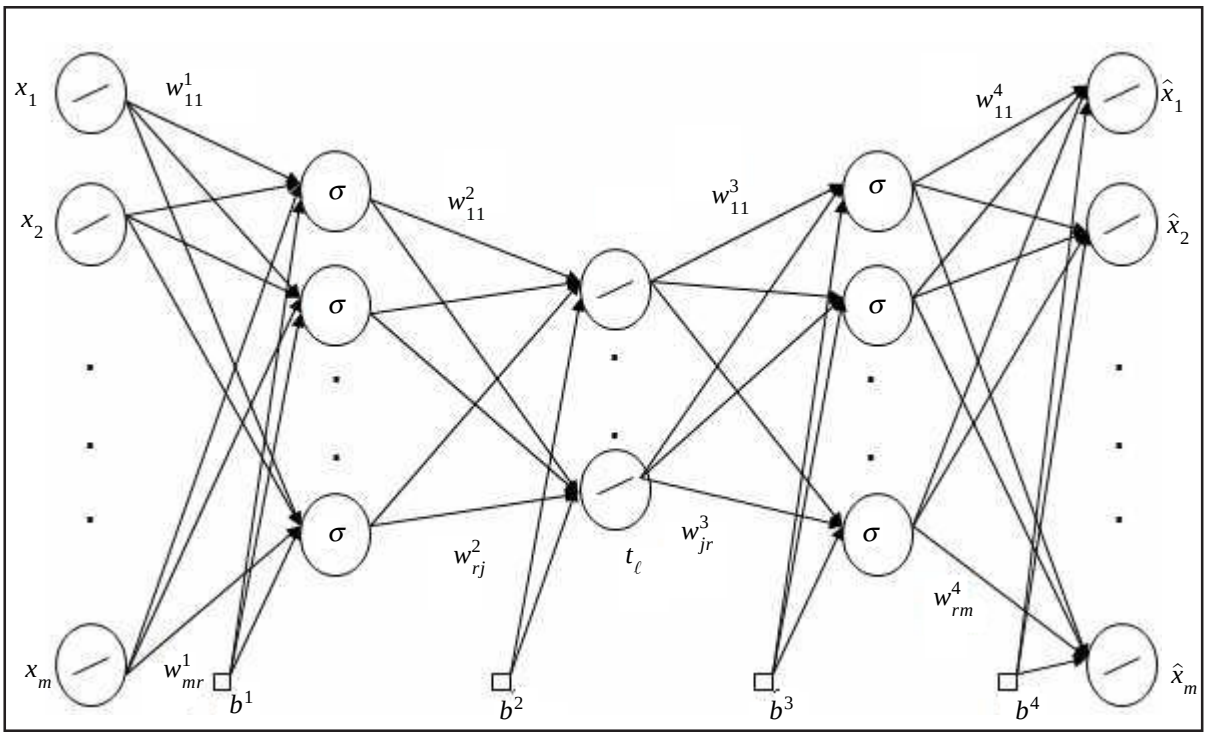

Figure 3. Neural Network

$$SPE = \| \tilde{x} \|^2 = x^T \tilde{P} \tilde{P}^T x = x^T \tilde{C} x \tag{6}$$

with $e = (I - \tilde{P}\tilde{P}^T)\, x$ present the residue of the vector x.

We consider a vector of measures x which follows a normal distribution $N(0, \Sigma)$ with λ_i, $i = 1, ..., m$ are the eigenvalues of Σ. The *SPE* limit detection is determinate by Jackson and Mudholkar [22] as follow:

$$C_\alpha = \theta_1 \left[\frac{1 - \theta_2 h_0 (1 - h_0)}{\theta_1^2} + z_\alpha \frac{(2\theta_2 h_0^2)^{1/2}}{\theta_1} \right]^{1/h_0} \tag{7}$$

with

$$\theta_1 = \sum_{i=1}^{m} \lambda_i,\ \theta_2 = \sum_{i=1}^{m} \lambda_i^2,\ \theta_3 = \sum_{i=1}^{m} \lambda_i^3,\ h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2} \tag{8}$$

3.4 Defect localization

The calculation of contribution to a variable by the SPE method is determined by considering the large square residue associated with a single variable:

$$SPE_i = \tilde{x}_i^2 \tag{9}$$

Statistical flaw detection has the quadratic form defined as follows:

$$statistic\,(x) = x^T B x = \| x \|_B^2 \tag{10}$$

Whith B is equivalent to $\tilde{C}$ in case of SPE method use and D for Hotelling T^2 method use.

Statistic (x) is given by the following equation:

$$statistic\,(x) = x^T Bx = \| B^{1/2} \|^2 = \sum_{i=1}^{N} (\xi_i^T B^{1/2} x)^2 = \sum_{i=1}^{N} c^{statistic} \qquad (11)$$

The contribution of x_i is appointed by $c_i^{statistic} = (\xi_i^T B^{1/2} x)^2$, and ξ_i is the i^{th} column of the identity matrix and direction x_i.

For statistical SPE, the contribution calculation for a variable is defined in [91] as follows:

$$c^{SPE} = (\xi_i^T \tilde{C}^{1/2} x)^2 = \tilde{x}_i^2$$

4. Validation of Proposed Approach

4.1 Algorithm

The steps of the proposed method of ECG defect detection are described is the following diagram:

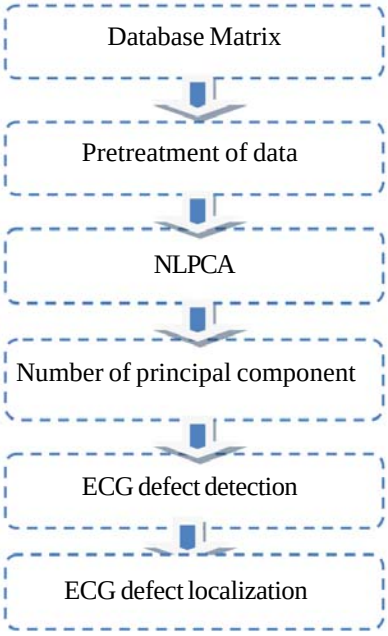

Figure 4. Proposed algorithm

The first step is to prepare the data matrix using the characteristic waves and segments of the ECG.

Secondly, a pretreatment is called to refine the data matrix by reducing its dimensions and center its measures.

In the third step, we apply the principal component analysis based on neural network. We introduce a multilayer's neural network. Several learning algorithms are used and compared in terms of performance.

The number of principal components to retain is then determined. Indeed, there are several methods for this fact telque: the cumulative percentage of total variance, average eigen values .... We chose the method based on the criterion of Webb since it is the most compatible with the NLPCA.

In the detection phase defects, we compared the results of two techniques: SPE and Hotteling methods. The first method gives more effective results and will be introduced for the detection of defects.

Once detected, we are called to localize defects to determine their origin. The calculation of contributions is introduced to determinate the defectives variables.

4.2 Database matrix

The database MIT/BIH provided by laboratory of the hospital of Beth Israel on Bosten in collaboration with the institute MIT is formed by different bases. This database contain 25 registrations left are chosen among the same group to include arrhythmias less common but medically significant. The registrations are sampled on 360 HZ per canal with the resolution of 11 bits on the range 10 mv. Two cardiologists independently annotate each ECG.

The matrix of introduced data is 9 variables and 529 measures (9×529) as shown in figure 4. These variables are defined as follow:

- *PA*: indicate the amplitude of the wave *P*.
- *RA*: it is the amplitude of the wave *R*.
- *QA*: it is the amplitude of the wave *Q*.
- *TA*: indicate the amplitude of the wave *T*.
- *SA*: defined the amplitude of the wave *S*.
- *QS*: it is the interval which separates the waves *Q* and *S*.
- *QP*: it is the interval which separates the waves *Q* and *P*.
- *RR*: it is the segment separating two successive *R* waves.
- *ST*: it is the segment separating the waves S and *T*.

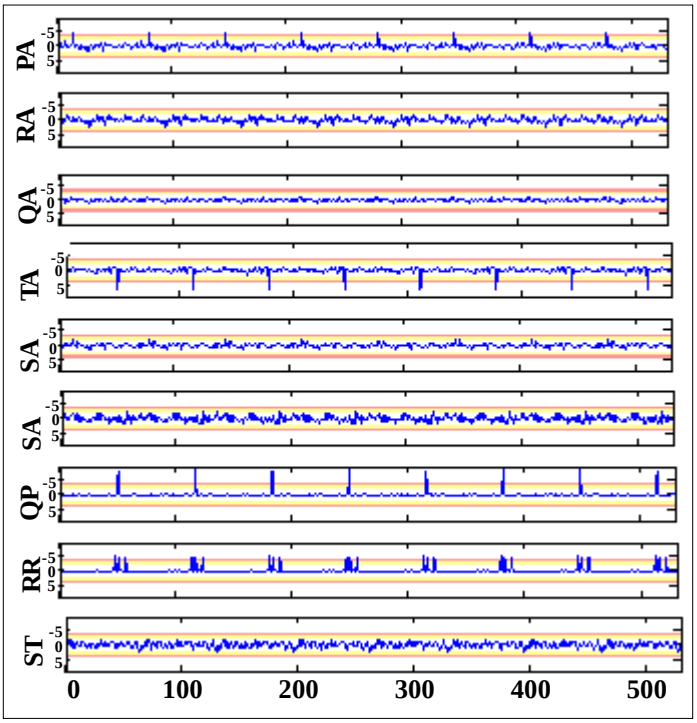

Figure 5. Evolution of variables

4.3 Pre Treatment

To find successful results, we are called to reduce and to center the matrix of studied data. The figure below presents the various variables after this operation:

4.4 Results

We begin with the determination of the number of principal components by using the index $\varepsilon(\ell)$. The index $\varepsilon(\ell)$, calculated for the various variables, is represented in the following table:

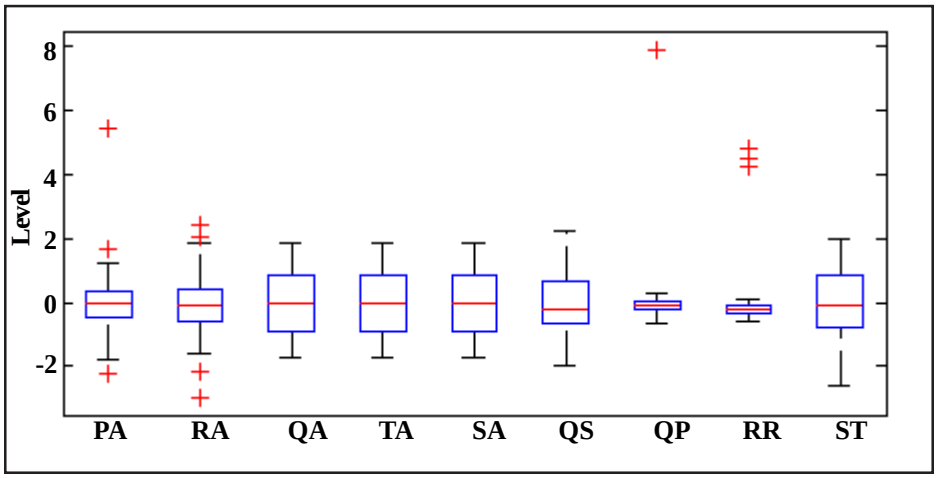

Figure 6. Reduced and centrated variables

ℓ	1	2	3	4	5	6	7	8	9
$\varepsilon(\ell)$	1.00	0.87	0.88	0.86	0.81	1.13	0.95	1.10	1.23

Table 1. Index $\varepsilon(\ell)$ for Each Variable

So the number of principal components to be retained in the model ACPNL is $\ell = 4$. We have four neurons in the layer in the middle where the index $\varepsilon(\ell)$ is minimal.

These fixations are proposed in the toolbox nntraintool [23] and defined as follow:

• **Learning:** 70 % of data are presented to the network during the learning and the network is adjusted according to its error (the curve in blue).

• **Validation:** 15 % of data are used to measure the generalization of network and interrupt the learning when the generalization stops improving (the curve in green).

• **Test:** 15 % of data have no effect on the learning and supply a test independent from performance of hanging network and after the learning (the curve in red).

For the validation of our approach and during the phase of learning of the network we used the learning of back-propagation and its various optimizations. Among these algorithms:

• Gradient descent with momentum and adaptive learning rate back- propagation (traingdx).

• Gradient descent back- propagation (traingd).

• Gradient descent with momentum back propagation (traingdm).

• Conjugate gradient back-propagation with Polak-Ribiére updates (traincgp).

• Scaled conjugate gradient back-propagation (scg).

In the following table we represent a comparative study of the different algorithms of the auto-associative learning:

The previous table shows that the algorithm based on the Gradient descent with momentum and adaptive learning rate back-propagation have the closest value of mean square error to the threshold that is equal to 0.1 (figure 7). That is why this algorithm will be included in rest of our study.

Figure 7 shows the mean squared error (mse) and we note that the best validation performance is equal to 0.2 at iteration 171. This value is the closest one to the goal which is equal to 0.1.

	traingdm	traingd	trainscg	traincgp	traingdx
Performance (mse)	1.4	1.5	0.4	0.8	0.2
Time	89 sec	90 sec	6 sec	4 sec	20 sec
Iterations	1000	1000	40	18	171
Structure	8-12-2-12-8	8-12-2-12-8	8-10-2-10-8	8-8-1-8-8	8-8-1-8-8

Table 2. Index $\varepsilon(\ell)$ for Each Variable

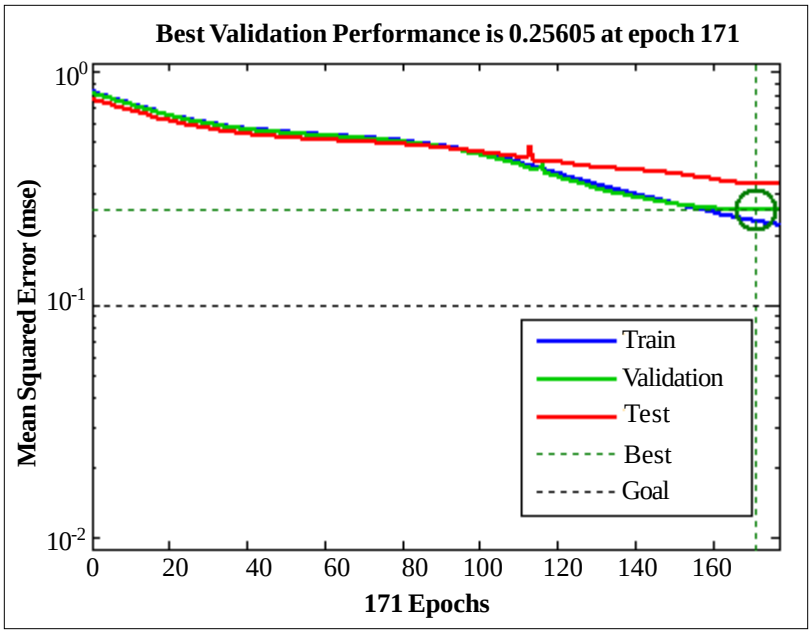

Figure 7. Best performance

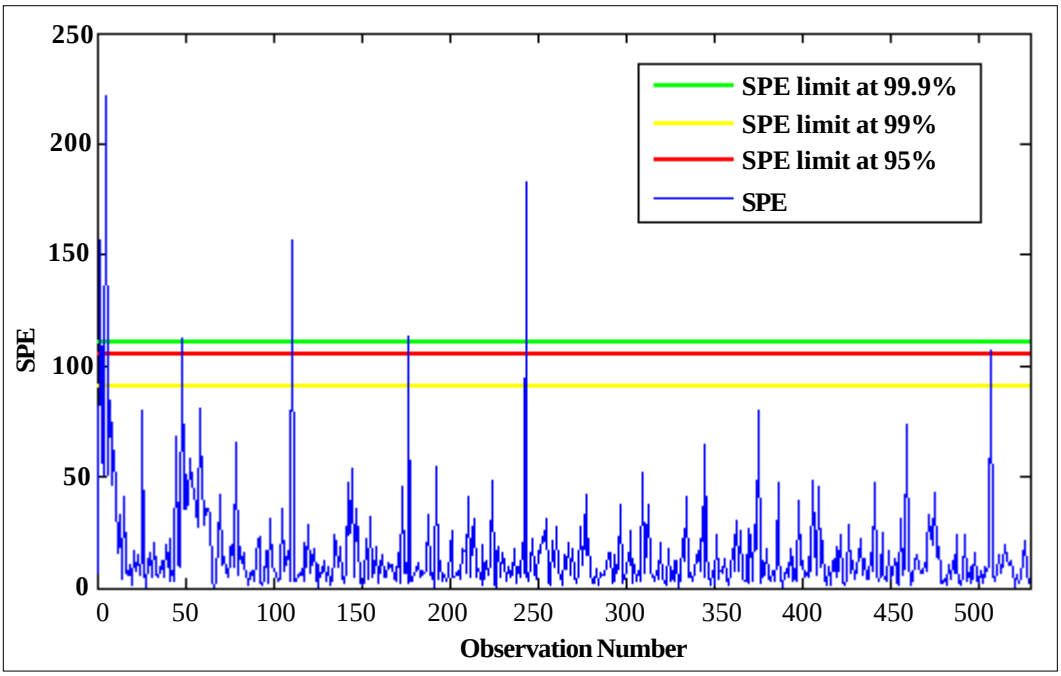

Figure 8. Defect detection using SPE method

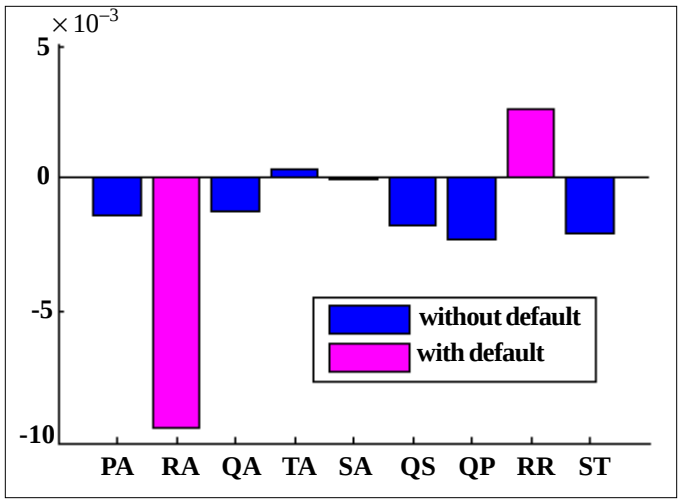

Figure 9. Defect localization

For the defect detection we use the SPE method defined in III.B. The following figure shows a defect detection using the SPE statistic.

In Figure 8 we use three control limits for the SPE statistic, at 99.9% (green), 99% (red) and 95% (yellow). Exceeding these three limitations proves the existence of a default .

The following figure shows the differents contribution of variables, the variables RA and RR which have the highest contributions are considered defected.

5. Conclusion

In this paper, a method for the detection of defects in an electrocardiogram is approached basing on the non linear principal components analysis. The proposed approach introduces a network of auto-associative neuron in five layers. The matrix of data is established by 9 variables determinate from the waves and characteristic segments of the studied ECG. First, we make a preprocessing which consists in reducing and in centering the matrix of data what is going to increase the reliability of the results. The proposed NLPCA is applied by using the algorithm of learning based on the Gradient descent with momentum and adaptive learning rate back- propagation. This algorithm is chosen after several tries and showed better results. The performance in the iteration 171 found equal to 0.25 is the closest value to the threshold (0.1). In the detection of the defects we use the quadratic error SPE and we apply it to a sick ECG. To define exactly the location of the detected defect we calculate contributions of all variables. RA and RR which have the highest contributions are considered defected variables.

References

[1] Algra, A., Zeelenberg, H. L. B. C. (1987). An algorithm for computer measurement of QT intervals in 24 hour ECG. *In*: Computer in Cardiology. Los Alamitos, CA, *IEE computer Society Press*, p. 117-119.

[2] Bahoura, M., Hassani, M., Hubin, M. (1997). DSP implementation of wavelets transform for real time ECG waveforms detection and heart rate analysis. *Comput.Meth.Programs Biomed*, (52) 35-44.

[3] Clavier, L., Boucher, J. M. (1996). Segmentation of electrocardiograms using a Hidden Markov Model. *In*: 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society October 31-November 3.

[4] Dokur, Z., Olmez, T., Yazgan, E., Ersoy, O. (1997). Detection of ECG waveswaveforms by neural networks. *Med.Eng.Phys*, 19 (8) 738-741.

[5] Chaouch, H., Ouni, K., Nabli, L. (2012). Segmenting and supervising ECG signal by combining CWT & PCA, *IJCSI International Journal of Computer Science Issues*, 9 (2) 1, March.

[6] Chaouch, H., Ouni, K., Nabli, L. (2001). ECG signal monitoring using linear PCA, *International Journal of Computer Applications* (0975 – 8887), 33 (8), November.

[7] Kaiser, H. F. (1960). The application of electronic computers to factor analysis, *Educational and Psychological Measurement*, 20, p. 145-151.

[8] Giancarlo et, D., Chiara, T. (2002). Cross-validation methos in principal component analysis: a comparision, *Statistical Method and Applications*, 11, p.71-82.

[9] Xu et, G., Kailath, T. (1994). Fast estimation of principal eigenspace using lanczos algorithm, *SIAM Journal on Matrix Analysis and Applications*, 15 (3) 975-994.

[10] Nabli, L., Ouni, K. (2008). The indirect supervision of a system of production by the principal components analysis and the average dynamics of the metrics, *International Review of Automatic Control* (IREACO), p. 560-567, November.

[11] Nabli, L., Ouni, K., Messaoud, H. (2008). Méthode de surveillance indirecte d'un système de production par l'analyse en composantes principales, La 5ème conférence internationale JTEA 2008, *Journées Tunisiennes de l'Electrotechnique et de l'Automatique*, 2-4 Mai.

[12] Nabli, L., Ouni, K., Haj, H. (2008). Salem. Approche Multi agents pour la surveillance indirecte d'un système de production par l'analyse en composantes principales, la Conférence Internationale Francophone d'Automatique CIFA 2008, p. 128-136, 3-5 Septembre.

[13] Pata, D., Kumardas, M., Pradhan, S. (2010). Integration of FCM, PCA and Neural Networks for classification of ECG arrhythmias, *IJCS* 36-3-05.

[14] Harakat, M. F., Mourot, G., Ragot, J. (2005). Détection de défauts de capteurs d'un réseau de surveillance de la qualité de l'air, *Journal Européen des Systèmes Automatisés*, *JESA*, p. 417-436, 2005.

[15] Besse et P., Ferré, L. (1993). Sur l'usage de la validation croisée en analyse en composantes principales. Reveue de statistique appliquée, *XLI* (1), p.71-76.

[16] Sylvester, J. J. (1889). On the reduction of a bilinear quantic of the nth order to the form of a sum of *n* products by a double orthogonal substitution, *Messenger of Mathematics*, 19, p. 42-46.

[17] Kramer, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks, *AIChE Journal*, 37 (2) 233-243.

[18] Dunia et, R., Qin, S. J. (1998). A subspace approach to multidimensional fault identification and reconstruction, *AIChE Journal*, 44, p. 1813-1831.

[19] Hastie et, T., Stuetzle, W. (1989). Principal curves, *Journal of the American Statistical Association*, 84 (406) 502-516.

[20] Chauvin et, M., Leenhardt, A. (2002). Electrocardiogramme: indications et interprétations, Hermes.

[21] Webb, A. R. (1999). A loss function to model selection in nonlinear principal components, *Neural Networks*, 12, p. 339-345.

[22] Jackson et, J. E., Mudeholkar, G. S. (1979). Control procedures for residuals associated with principal component analysis, *Technometrics*, 21 (3) 341-349.

[23] Demuth, H., Hagan et, M. T., Beale, M. (2010). Neural Network Toolbox, User Guide, Version 7.