Enhanced Multivariate Process Monitoring for Biological Wastewater Treatment Plants

Alima Chaouch, Khaled Bouzenad, and Messaoud Ramdani
Laboratory of Automatic and Signals-Annaba (LASA)
Department of Electronics, Badji-Mokhtar University, Annaba, BP. 12, 23000, Annaba, Algeria
E-mail: auto_loulou@hotmail.fr, khaled_bzd@yahoo.fr, messaoud.ramdani@univ-annaba.org

***Abstract*—Multivariate statistical methods for the analysis, monitoring and diagnosis of process operating performance are becoming more important because of the availability of on-line process computers which collect measurements on large numbers of process variables. Principal component analysis (PCA) has been used successfully as a MSPC tool for detecting faults by extracting feature information from complex data. However the traditional linear PCA is not well adapted to complicated nonlinear systems; therefore Non-linear principal component analysis (NLPCA) is a non-linear generalization of standard linear PCA. NLPCA can be achieved by using a neural network with an auto-associative architecture. In order to monitor process performance, the SPE index is used for anomalies detection step. Then the fault localization by exploiting reconstruction method as an alternative to contribution plots approach. The obtained results on real data demonstrate the technique effectiveness applied for monitoring wastewater treatment plant.**

***Index Terms*—process monitoring, multivariate statistical process control, diagnosis, wastewater treatment plant, auto-associative neural network, nonlinear principal component analysis**

I. Introduction

Protecting the environment is a necessity; or an emergency even to limit the damage caused by different types of pollution. Among the most polluting factors, wastewater discharged from different origins in nature, in the wadis and the sea, have already caused a lot of damage both human and material. Hence, stricter standards for the operation of wastewater treatment plants (WWTP) have been imposed in order to generate a limit to the quality of toxic and organic matter released in industrial and municipal effluent. The wastewater treatment can separate water pollutants with different physical, chemical or biological processes. The goal of these treatments is to reduce the maximum amount of pollutants in the wastewater. The water returned to the environment must be equipped with sufficient quality to not degrade the receiving environment. For local stations as for industrial sites, the wastewater treatment is an elaborate process that usually involves various physicochemical and/or biological techniques. Biological technique is an essential operation for wastewater treatment, where the main objective is the elimination of the organic pollutant compounds, degradation of the organic components is assured by microorganisms (bacteria) that consume organic matter in the presence of oxygen (aerobic method) or without oxygen (anoxic method). Inside a biological wastewater treatment plant, the activated sludge process is the most widely used biological treatment of wastewaters containing carbon and nitrogen pollutants [1], were the most successful model is the activated sludge process Model No.1 [2]. This model trigged the general acceptance of the biological modeling were default stoichiometric and kinetic parameters have been proposed and proved to give realistic results. The use of efficient models in controller design is important. However, demand of efficient monitoring techniques is high in order to increase the efficiency of the system. Large number of process monitoring including fault detection and diagnosis based on statistical process control has increased, almost exponentially over the last few decades The wastewaters should not be directly rejected into the natural environment, because they can generate serious environmental and public health problems. The treatment or the wastewater purification aims thus to reduce the polluting load which they convey. Consequently they should be directed towards sewage treatment plants whose role is to concentrate the contained pollution in the form of the residues with small volume (sludge) and to reject a purified water respondent to specific norms, that thanks to physic-chemical or biological process, From the entrance of the station to discharge into the natural environment, the different stages of wastewater treatment are:

Firstly, pretreatment aims to eliminate the coarser solid or particulate elements, voluminous waste (screening), sands (desanding) and fatty substances (Degreasing). Secondly, primary treatment so as to expel the dissolved pollutant load and remaining suspension materials, physic-chemical treatments allow agglomerating these particles by the addition of coagulants and flocculation agents.

Finally, the secondary treatments recovers the elimination techniques of soluble polluting load (carbon, nitrogen and phosphorus), treatments are based on

Manuscript received September 7, 2013; revised November 12, 2013.

doi: 10.12720/ijoee.2.2.131-137

biological processes nature for instance the activated sludge process. The activated sludge is a biomass (bacteria) in suspension in stirred basins, the bacteria find favorable conditions for their development (of oxygen and the carbonaceous substrates) [3]. The treatment of activated sludge comprises two stages: Nitrification: It takes place in aired zone, the organic nitrogen and ammoniacal contents in water are oxidized out of nitrite via bacteria, the nitrites (NO_2^-) are then oxidized out of nitrates (NO_3^-).

Denitrification: It takes place in anoxic upstream area and channels during the anoxic phases. De-nitrification brings heterotrophic bacteria that use oxygen bound to nitrogen molecules, it results a clearance of gaseous nitrogen *(N_2)*. Finally, the clarification step in the clarifier ensures the separation between sludge and treated water. In our case of study the monitoring approaches presented in this document has been tested by using real data from the Annaba WWTP, situated in 450 kilometers from Algiers and eight (km) from the east of Annaba on the municipality of El Bouni. The purification process used is activated sludge and includes two treatment sectors, one for water and the other for muds. The medium flow to purify is of 83,620 m^3/j and peak flow during dry weather 5923 m^3/h (see Fig. 1).

Figure 1. Water cycle in the wastewater treatment plant of Annaba.

The data analysis of wastewater process is based on multivariate statistical methods for monitoring and fault sensors diagnosis through a nonlinear projection technique NLPCA based on neural network. This is shown in the following, using a real database collected from the wastewater treatment plant of Annaba during one time of three month of year (2011-2012). The variables which took account in this study are the following ones: the wastewater $Q_{inf}(m^3/j)$, wastewaters temperature $T_{\mathrm{inf}}(C°)$, purified water temperature $T_{\mathrm{eff}}(C°)$, PH_{inf}, PH_{eff}, dissolved oxygen input $O_{2in}(mg/l)$, dissolved oxygen of treated water $NO_{2eff}(mg/l)$, Conductivity input $C_{inf}(\mu s/cm)$, Conductivity output $C_{eff}(\mu s/cm)$, ammoniacal nitrogen $NH_{4eff}(mg/l)$, nitrates $NO_{3eff}(mg/l)$, and Nitrites $NO_{2eff}(mg/l)$. The fault sensor detection technique adopted in this work is based on detection indicator *SPE*, that which defines in residual space. To locate the defective sensor, the calculating contribution approach to the detection index *SPE* is exploited. In order to mitigate the gaps of the traditional contribution method, the reconstruction approach by NLPCA, PCA is employed. This alternative was based on the elimination of defects influence on the detection index *SPE* by reconstruction offending variable using NLPCA neuronal model [4] and [5].

II. Nonlinear Principal Component Analysis

The increasing complexity of automated systems is accompanied by the ever-stronger requirements in terms of performance constraints, these requirements place to develop more adequate control systems approaches. In this context, the multivariate statistical process control approaches are used to analyze all the information provided by the sensors useful for the command automated systems. These approaches are based on linear projection techniques such as principal component analysis *PCA* which is used as a tool for modeling linear correlations between variables process. However, in the case of nonlinear systems, the *PCA* is unable to reveal the structures or non-linear relationships, which requires the use of other approaches. Thus, the extension of *PCA* to treat nonlinear problems was developed. The global model of the *PCA*, whether linear or nonlinear is composed of two sub-models, a sub-model generates a projection data of dimension n to a the main components subspace of dimension ℓ and the second sub-model is the decompression operation where he plans principal components $\mathfrak{R}^{\ell}$ to the original data $\mathfrak{R}^{n}$. In the linear case, the first model in the *PCA* includes the eigenvectors matrix of the correlation matrix data $\hat{P}$ and the second sub-model is expressed by the inverse matrix $p^{-1}=p^{T}$, note that the two sub-models are orthogonal. On the other hand, in the nonlinear case, the projection matrices are replaced by non-linear functions ψ and φ (see Fig. 2).

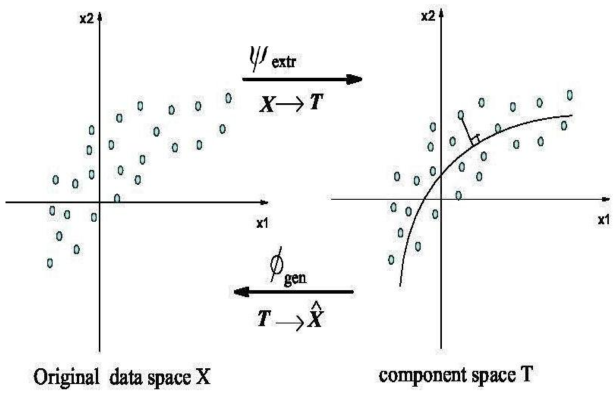

Figure 2. Nonlinear reduction of dimensionality data.

when we reduced the size of the data $X\in\mathfrak{R}^{n}$ for non-linear principal components $T\in\mathfrak{R}^{n}$ the non-linear projection function ψ: $\mathfrak{R}^{n} \rightarrow \mathfrak{R}^{\ell}$ is defined by [6].

$$T=\psi\left(X\right) \tag{1}$$

The non-linear function φ: $\Re^{\ell} \rightarrow \Re^{n}$ is the reconstruction function or data generation, it is defined by [6].

$$\widehat{X} = \varphi(T) \quad (2)$$

It is concluded that the data matrix X can be expressed by the reconstruction matrix $\widehat{X}$ by adding the estimation error (residual matrix) $\tilde{x}$:

$$X = \widehat{X} + \tilde{X} = \varphi(T) + E \quad (3)$$

III. The NLPCA Model Based on Artificial Neural Networks

Kramer (1991) proposed a non-linear *PCA* extension using a neural network with a self-associative topology includes five layers with a throttle layer (bottleneck layer). Thus, the neural network is a multi-layer perceptron (*MLP*) contains three layers between the input and output variables (see Fig. 3).

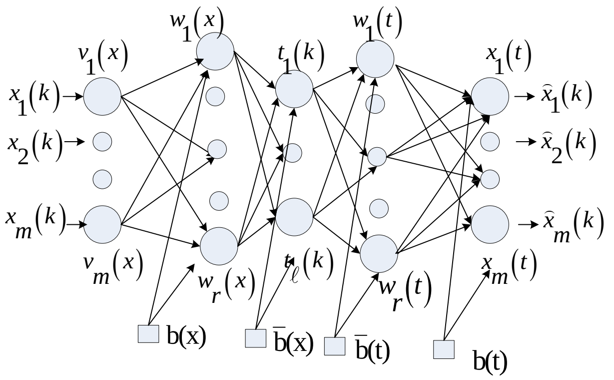

Figure 3. Auto-associative neuronal network of *NLPCA* proposed by Kramer.

A transfer function Ξ_1 performs a projection of the input vector X of dimension m to the first hidden layer (coding layer), expressed by $h^{(t)}_{j\in(1\ldots r)}$, r is the number of neurons in the first hidden layer:

$$h_j^{(x)} = \Xi_1((V^{(x)})X + b^{(x)}) = \Xi_1(\sum_{i=1}^{m} v_{ij}^{x} + b_j^{x}) \quad (4)$$

$b^{(x)}$= The bias includes r parameters.

$V^{(x)} \in \Re^{(m\times r)}$: represents the weight matrix.

The second transfer function Ξ_2 projects the data of the first hidden layer to the throttling layer (bottleneck layer) containing less neurons, which represent the non-linear principal component t_ℓ =(1...ℓ), ℓ is the number of neurons in restrictor layer [7]-[10].

$$t_i = \Xi_2(w^{(x)}h^{(x)} + \bar{b}^{(x)}) = \Xi_2(\sum_{j=1}^{r} w_{ji}h_j^{(x)} + \bar{b}_i) \quad (5)$$

Then a non-linear function Ξ_3 projects main components, t_ℓ to the last hidden layer (decoding layer) is the number of neurons in the $h^{(t)}_{j\in(1\ldots r)}$, r is the number of neurons in the third hidden layer:

$$h_j^{(t)} = \Xi_3(w^{(t)}t_\ell + b^{(t)})_j) = \Xi_3(\sum_{i=1}^{\ell}(w_{ij}^{t}t_i + b_j^{(t)}) \quad (6)$$

Finally there is the last identity function Ξ_4 which projects the data in the hidden layer $h^{(t)}$ to the output vector $\hat{x}$.

$$\hat{x}_i = \Xi_4(v^{(t)}h^{(t)} + \bar{b}^{(t)}) = \Xi_4(\sum_{j=1}^{r} v_{ji}^{(t)}h_j^{(t)} + \bar{b}_i^{(t)}) \quad (7)$$

Generally the network uses a supervised learning method, it is considered an optimization algorithm, where this is done iteratively by changing the weights according to the gradient of the cost function (see equation (8)) and the gradient is minimized by the back propagation method [10] and [11].

$$E(k) = \frac{1}{2}\sum_{i=1}^{m}(x(k) - \hat{x}(k))^2 \quad (8)$$

IV. Detection and Localization by NLPCA

A. Fault Detection

The Most conventional methods of Statistical process control (SPC) provides control charts interpretable only a reduced number of variables. In addition, univariate charts provide quantitative information by ignoring the effect of the correlation between variables. For this purpose several multivariate extensions of univariate approaches (Shewhart charts, CUSUM and EWMA) have been proposed in literature. The first multivariate charts have been developed represent the χ^2 and the Hotelling T^2 maps. In this context, the typical index used for the detection of abnormal operation is *SPE* index conjunction with the *NLPCA* neuronal model. The *SPE* index provides fault detection in the residual subspace [11]. Its expression at the instant k is given by:

$$SPE(k) = e(k)\, e(k)^T \quad (9)$$

where the vector of estimation errors has the following formula:

$$e(k) = \tilde{x}(k) = x(k) - \hat{x}(k) \quad (10)$$

where the estimated data at the instant k is represented by:

$$\hat{x}(k) = \varphi(t(k)) = \varphi(\phi(x(k))) \quad (11)$$

With such an indicator, the process is considered in normal operation to the k^{th} observation if:

$$SPE(k) < \delta_\alpha^2 \quad (12)$$

$$\delta_{\alpha}^{2} = g\chi_{h,(1-\alpha)}^{2} \tag{13}$$

δ_{α}^{2}: Represents a detection threshold of *SPE*, *h* is the degree of freedom, the coefficient *g* can be estimated by the average *m* and the variance *v*:

$$g = \frac{v}{2m} \tag{14}$$

$$h = \frac{2m^2}{v} \tag{15}$$

$$\delta_{\alpha}^{2} = \frac{v}{2m}\chi_{\frac{2m^2}{v},(1-\alpha)}^{2} \tag{16}$$

α: is the predefined significance level.

B. Fault Localization

- Localization by calculation of the contributions

The most conventional and widely used in the context of *NLPCA* for fault location approaches are based on the calculation of contributions. The principle of these methods is to calculate the contributions of the different variables to the indicators used for the detection of defects. Thus, the contributions principle is generally based on the share quantification of each variable in the calculation of the detection index *SPE*. The contribution $Cont_j^{SPE}(k)$ of j^{th} variable at the time *k* is defined by the equation (17):

$$Cont_j^{SPE}(k) = [\tilde{x}_j(k)]^2 = [x_j(k) - \hat{x}_j(k)]^2 \tag{17}$$

where $x_j(k)$ is the j^{th} element of the measurement vector *x* at the instant *k*. The localization of defects based on the contributions shows that those guarantee a correct diagnosis only if the faults are simples and have a large amplitudes. Otherwise, contributions approaches can generally consider another variable is faulty. Thus, it is difficult to isolate those really in default. Moreover, the contributions do not allow insulation of the multiple failures where several variables are simultaneously at fault because of the correlation between the variables. This correlation was the key of a better decisive diagnosis based on reconstruction approach.

- Nonlinear reconstruction

The principle of this approach is applied when it detects that a sensor is faulty, it consists to reconstruct the value of its measurement based on other variables. The reconstruction designated by $\hat{x}$ can be obtained by an iterative algorithm as indicated on the Fig. 4 The estimated value $x_i(k)$ is re-estimated (re-injected) by *NLPCA* until convergence as follows:

$$\hat{x}_i = \zeta_i^T \varphi\left(\psi\left(x_i\right)\right) \tag{18}$$

where $x_i = [x_1, x_2, \ldots, \hat{x}_i, \ldots, x_m]^T$ and ξ*i* the i^{th} column of the identity matrix. As the algorithm is iterative, we must choose initialization value for *xi*, for this it was chosen $x_i^{(0)} = x_i$, the algorithm converges in a few iterations because the convergence is very fast.

The localization is performed by calculating the detection indicator $SPE_i(k)$ after reconstruction the all variables. It is noted that the reconstruction the faulty variable eliminates the effect of the fault, thus this detection indicator $SPE_i(k)$ of the offending variable will be bottom its detection threshold $SPE_i(k) < \delta_{\alpha}^{2}$, while the other indicators of other variables are all above their respective thresholds.

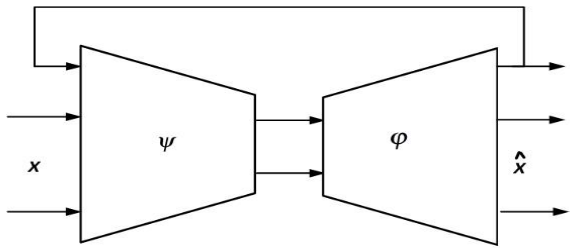

Figure 4. The reconstruction principle.

- Sensors Validity index (SVI)

After the detection of the presence of a fault, the localization is carried out by the comparison of the detection index before and after the reconstruction. This comparison is called sensor validity index $\eta_j^2(k)$:

$$\eta_j^2(k) = \frac{SPE_i(k)}{SPE(k)} \tag{19}$$

The validity index of a faulty sensor must converge towards zero.

V. Application to a Wastewater Treatment Plant of Annaba

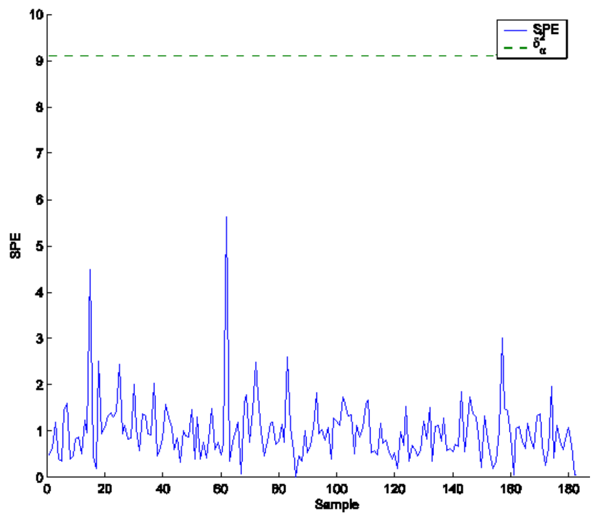

Figure 5. Statistical *SPE* to α = 0.05.

In order to illustrate the theoretical study of this work, we consider the simulation of an example process of biological wastewater treatment, which the data are real, collected from the wastewater treatment plant of Annaba in a period of three months of the year (2011-2012). A data matrix X has been consists of N = 180 observations which represent a normal operation of the process, have been reserved for the construction of *NLPCA* model. In particular, the data of such a matrix are centered and scaled using the means and standard deviations of reserved data to the model. For the monitoring model one selected a vector of measurements construct of 12 variables mentioned above.

Fig. 5 represents the statistical *SPE* of the data set during normal operation process to a confidence level for 95%.

After the construction of NLPCA model, and in order to illustrate a fault diagnosis using the different methods described in this article, an offset fault affecting the seven variable, SO_{2eff}, (oxygen at the output of WWTP), is simulated as a window that starts at sample k=45 and an amplitude belongs to the range of variation of the variable x_7. Fig. 6 shows the statistical *SPE*.

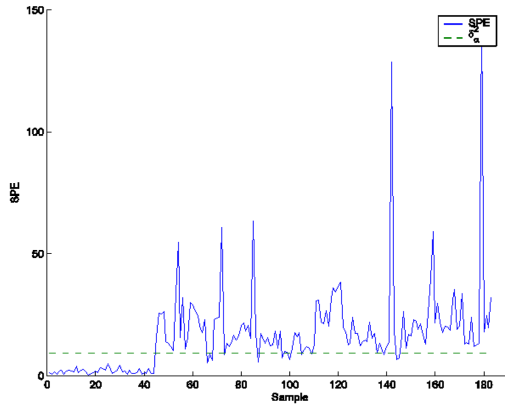

Figure 6. Statistical SPE to α = 0.05.

We note that the statistical SPE (see Fig. 7) immediately allows the detection of the fault. To identify the faulty sensor, it was exploited the contributions approach to detection index SPE.

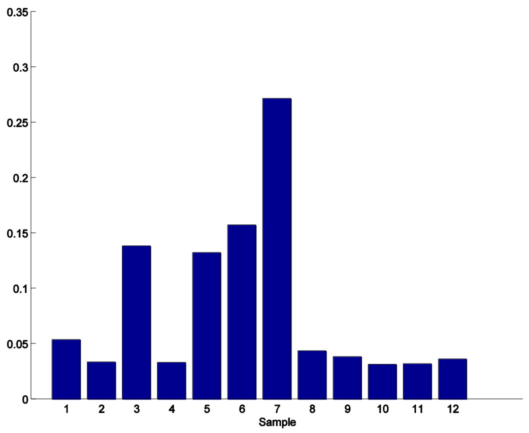

Figure 7. Localization of defect by the contributions to the SPE index.

Thus, according to the previous figure the variable x_7 having the largest contribution to SPE index calculated at time k = 45 of detection, this expresses that the seven variable, SO_{2eff}, (oxygen at the outlet of the WWTP) is failing. To isolate the infected variable, we used the reconstruction approach; the following figures illustrate the principle of the reconstruction method.

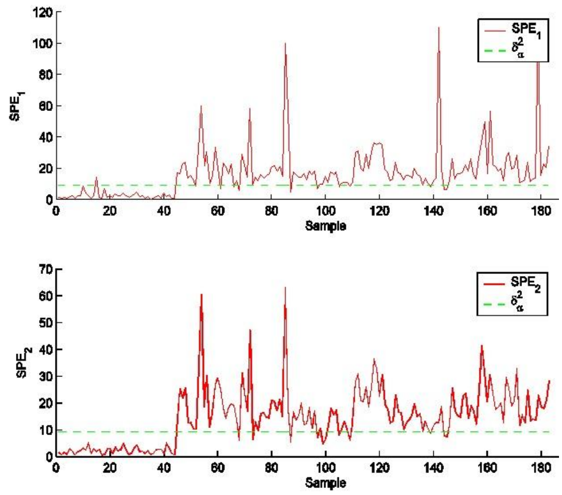

Figure 8. The statistics SPE1, SPE2 after the reconstruction of the variables x_1, x_2.

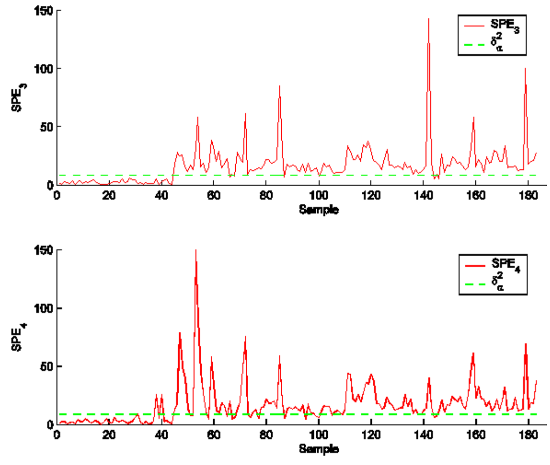

Figure 9. The statistics SPE3, SPE4 after the reconstruction of the variables x_3, x_4.

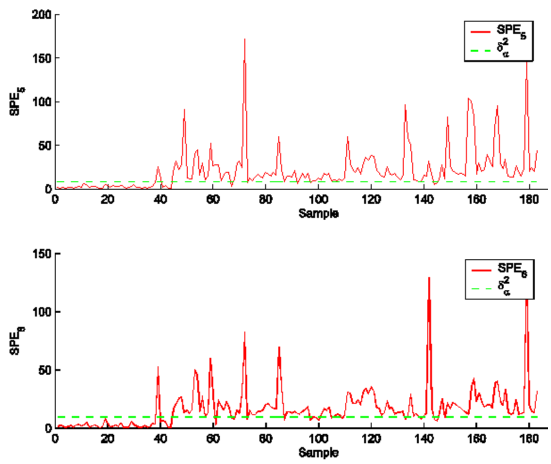

Figure 10. The statistics SPE5, SPE6 after the reconstruction of the variables x_5, x_6.

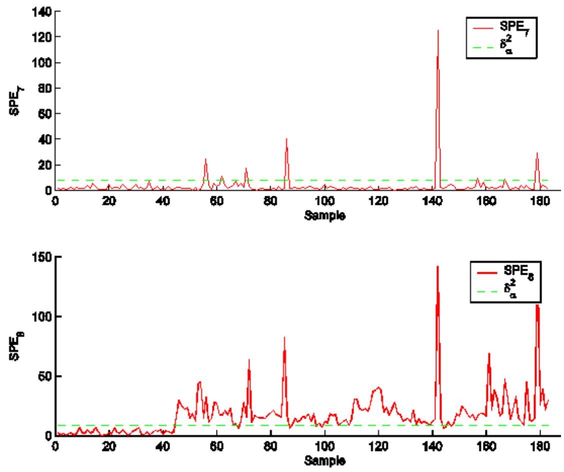

Figure 11. The statistics SPE7, SPE8 after the reconstruction of the variables x_7, x_8.

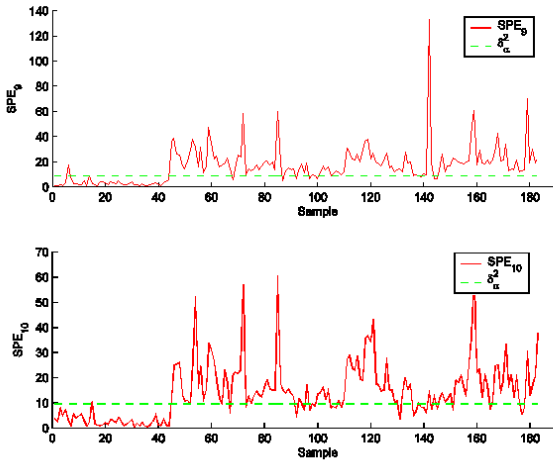

Figure 12. The statistics *SPE*9, *SPE*10 after the reconstruction. of the variables x_9, x_{10}.

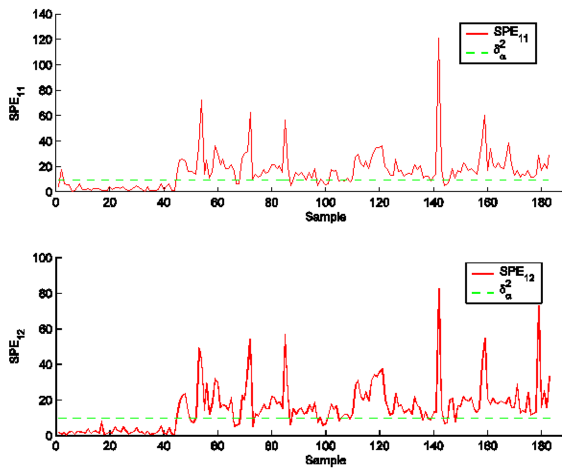

Figure 13. The statistics *SPE*11, *SPE*12 after the reconstruction. of the variables x_{11}, x_{12}.

According to the figures, Fig. 8, Fig. 9, Fig. 10, Fig. 11, Fig. 12 and Fig. 13, we see that all indicators SPE_i "$i \neq 7$" are above their thresholds, except statistical SPE_7 who corresponds to the x_7 variable is lower than its detection limit. It is also noted that the reconstruction of the x_7 variable eliminates the defect effect, this expresses that the x_7 sensor is infected. After the reconstruction the sensor validity index *SVI* is calculated to locate the fault. The Fig. 14 shows the isolation fault on the variables by *SVI* index. It is noted that the faulty sensor, his validity sensor index going down towards the zero.

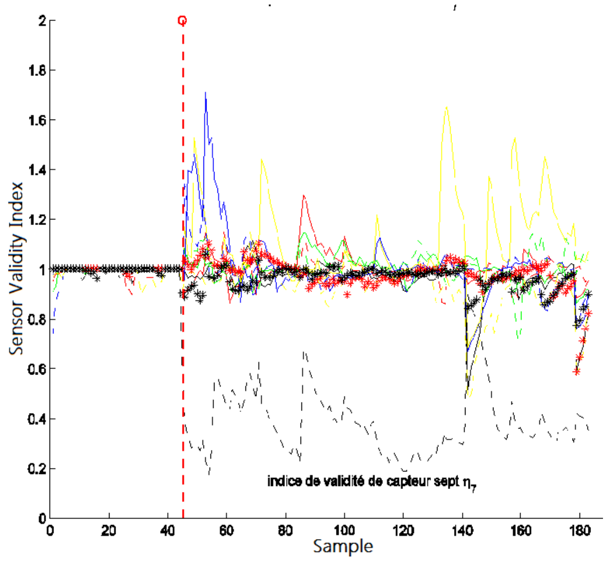

Figure 14. Sensor validity index with a fault in the seven variable at time k=45.

VI. Conclusion

The study described in this paper concerns the monitoring task of a nonlinear process of biological wastewater treatment plant, using *NLPCA* method of nonlinear multivariate statistical process control based on auto-associative neural network with a throttle layer. In the context of diagnosing and locating faults, affecting useful sensors to control the WWTP, Two localization approaches have been exposed: The calculation of the contributions to the detection index *SPE* and reconstruction approaches of the offending variable measures. This method has a better rate of correct failure diagnosis compared to the traditional contribution method. The results obtained in this paper are very encouraging in-depth study to apply other approaches proposed in the literature to best estimate the confidence limit and to improve the precision of the monitoring system.

References

[1] Y. M. Kim, H. U. Cho, D. S. Lee, D. Park, and J. M. Park, "Influence of operational parameters on nitrogen removal efficiency and microbial communities in a full-scale activated sludge process," *Water Research*, vol. 45, no. 17, pp. 5785-5795, 2011.

[2] M. Henze, C. P. L. Grady, W. Gujer, G. V. R. Maris, and T. Matsuo, *Activated Sludge Model No. 1: IAWPRC Science and Technical Report No. 1, I.A.W.Q*, U.K.: London, 1987.

[3] L. J. Zhao, T. Y. Chai, and Q. M. Cong, "Multivariate statistical modeling and monitoring of SBR wastewater treatment using double moving window MPCA," *Machine Learning and Cybernetics IEEE*, vol. 3, pp. 1371-1376, 2004.

[4] D. Aguado and C. Rosen, "Multivariate statistical monitoring of continuous wastewater treatment plants," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 7, pp. 1080-1091, 2008.

[5] H. Haimi, M. Mulas, F. Corona, and R. Vahala, "Data-derived soft-sensors for biological wastewater treatment plants: An overview," *Environmental Modelling & Software*, vol. 47, pp. 88-107, September 2013.

[6] J. M. Lee, C. Yoo, S. W. Choi, P. A. Vanrolleghem, and I. B. Lee, "Nonlinear process monitoring using kernel principal component analysis," *Chemical Engineering Science*, vol. 59, no. 1, pp. 223-234, 2004.

[7] C. K. Yoo, P. A. Vanrolleghem, and I. B. Lee, "Nonlinear modeling and adaptive monitoring with fuzzy and multivariate statistical methods in biological wastewater treatment plants," *Journal of Biotechnology*, vol. 105, no. 1, pp. 135-163, 2003.

[8] E. Oja and J. Karhunen, "On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix," *Journal of Mathematical Analysis and Applications*, vol. 106, no. 1, pp. 69-84, 1985.

[9] R. Dunia, S. J. Qin, T. F. Edgar, and T. J. McAvoy, "Use of principal component analysis for sensor fault identification," *Computers & Chemical Engineering*, vol. 20, pp. S713-S718. 1996.

[10] K. Bouzenad, M. Ramdani, N. Zermi, and K. Mendaci, "Use of NLPCA for sensors fault detection and localization applied at WTP," *Journal of E-Technology*, vol. 4, no. 3, pp. 138-148, August 2013.

[11] K. Mendaci, M. Ramdani, and T. Benzaraa, "Nonlinear multivariate statistical process monitoring of a water treatment plant," in *Proc. 5th International Conference on Modeling, Simulation and Applied Optimization, IEEE International Conference*, Hammamet: Tunisia, 2013, pp. 1-6.

Alima Chaouch was born in Algeria in 1984. Received her Engineer degree in electronic engineering option automatic in 2008 from Badji Mokhtar University of Annaba, She is currently a graduate student To get a master's degree at the 20 August 1955 University of Skikda. Her research interests are focused on Process Monitoring, Nonlinear principal component analysis, and Use of Hyperspaces for Fault Detection Isolation and Reconstruction.

Khaled Bouzenad was born in 1980. Received Magister's degree in 2009 and the engineer degree in telecommunications in 2006 from the University of Annaba. He is currently a PhD student at the University of Annaba, his research interests are focused on nonlinear principal component analysis, statistical process control, fault detection, identification, reconstruction and process monitoring.

Messaoud Ramdani received the doctorate degree in Automatic Control from the University of Annaba, Algeria, in 2006. He is currently a Lecturer in the department of Electronics, faculty of Engineering, University Badji-Mokhtar of Annaba, Algeria. He has published over 36 journal and conference papers. His research interests include pattern recognition, fuzzy logic, machine learning, data mining and statistical process control.