2013 Conference on Control and Fault-Tolerant Systems (SysTol)
October 9-11, 2013. Nice, France

Sensor Fault Detection, Localization and Reconstruction Applied at WWTP

K. BOUZENAD * M. RAMDANI ** A. CHAOUCH ***
Department of Electronics, Badji-Mokhtar University
BP. 12, 23000, Annaba, Algeria
E-mail: khaled.bouzenad@hotmail.fr *
messaoud.ramdani@univ-annaba.org **
auto_loulou@hotmail.fr ***

Abstract— **Principal Components Analysis (PCA) has been intensively studied and is widely applied in industrial process monitoring. The main purpose of using PCA is the dimensionality reduction by extraction of the feature space that still contain the most information in the original data set. Despite its success in this field, the most important obstacle faced is the sensitivity to noise, also the fact that the majority of collected data from industrial processes are normally contaminated by noise makes it unreliable in some cases. To overcome these limitations, several strategies have been used. One of these has been interested to combine the robustness theory with PCA method, such theory sonsists in robustifying the existing algorithms against noise or outliers. Fuzzy Robust Principal Components Analysis (FRPCA) is one of the result for such combination that acheive better result compared with the classical method. In this work the RFPCA method is used and compared with the classical one to monitoring a biological nitrogen removal process. The obtained results demonstrate the performances superiority of this method compared with the conventional one.**

Keywords: Diagnosis, Fault Detection And Localization, PCA, NLPCA, SPE: Squared Prediction Error, ANN: Auto-Associative Neural Network, SVI: Sensor Validity Index, WWTP: WasteWater Treatment Plant.

I. Introduction

In recent years the interest increased for the natural environment protection has made sure that the water treatment has become an important challenge. Indeed, the strict norms on discharges of wastewater treatment plants are becoming more and more stringent. The control of impacts on the natural environment and the operating costs, is not achieved only by an improvement of control installation type, for optimize the operation. Indeed, to operate correctly, this control system needs to know, at all times the process state. Any fault leads to generate a control signals which do not correspond to the process real state, where degradation of performance, reliability and sometimes a challenge to security and environmental quality, the operating diagnosis is an essential element of any process automation procedure. The objective of this paper is to apply a diagnosis procedure based on NLPCA on a dataset of sensors signals used in a WWTP control. At first time, and in the section (A): we apply the proposed approaches of diagnosis on the reduced ASM1 (reduced activated sludge model) using a reference model defined by the benchmark 624 (www.ensic.inpl-nancy.fr/COSTWWTP/Benchmark) [11], in the section (B), we present a second application on real data of annaba WWTP, located in eastern Algeria. The principal component analysis is then used to perform the sensor fault detection and localization applied at a first time, an reduced ASM1 of WWTP, followed an real dataset of annaba WasteWater Treatment Plant. In order to build a model we use a data matrix consisting of some "measures" generated by a simulator model proposed or by real sensors. The PCA is then a statistical technique which consists simultaneously to identify the relations between the variables of the process and to analyze and reduce the dimensionality of big size dataset. The PCA consists to replace a variables set by new variables uncorrelated two to two, of a smaller size and maximum variance, These new variables called principal components.

II. PRINCIPAL COMPONENT ANALYSIS

The Principal Component Analysis (PCA) is a method of data analysis family and more generally of multivariate statistics, which consists to transform a variables interrelated (called "correlated" in statistics) into new linearly uncorrelated variables with each others. These new variables are called "principal components" or axes. It allows the practitioner to reduce the information to limited number of components compared to the initial number of variables. It is a geometric approach (representation of variables in a new geometric space according to maximum inertia direction) and statistical approach (search independent axes explaining the most variability "variance" of data). Then, when we want to compress a set of random variables, the first axes of the PCA are a best choice, in terms of inertia or explained variance. We called principal axes the direction axes of the eigenvectors of the covariance matrix of the process variables, where the first axis that is associated with the biggest eigenvalue and the second axis orthogonal to the first, is associated with the second biggest eigenvalue...etc, and then the last axis is that associated with the smallest eigenvalue. The two or three first principal axes constitute the directions of the reduced space "Principal space" which belongs to the original data space. This approach is based on the projection of the original dataset on the new lower-dimensional space, and from the projection matrix, we can estimate our original information while minimizing the estimation error, in this sense the PCA can be considered as a minimization technique of the estimation error, otherwise, the estimated data must be approximately near to the original values.

III. NON LINEAR PRINCIPAL COMPONENT ANALYSIS

The principal component analysis has interesting properties for industrial processes monitoring. Unfortunately, in the industry, the most physical systems has a non-linear behavior and then the linearity property of the linear PCA, Figure (fig.2), pose always the problem of the inaptitude of this method to represent nonlinear data, because it is a linear projection and only the linear dependencies between the variables can be revealed. Hastie [1] proposes an approach for a generalization of PCA in the nonlinear case based on the principle of principal curves, this generalization is performed by a projection of data on curves instead of lines. Kramer [2] proposes the extension of the non-linear principal components analysis (NLPCA) using a neural network with five layers whose the weights are calculated by learning by minimizing the squared error between the inputs and outputs of network.

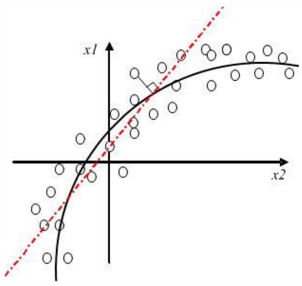


Fig. 1. Difference between linear and nonlinear PCA

This section is dedicated to non-linear extension of PCA (NLPCA) which allows to extract the linear and nonlinear relations between variables. The figure (Fig. 2) shows the principle of the general PCA, the overall model is composed of two sub-models: a sub-model of data compression projects the data of dimension m into principal components space of dimension ℓ and the second sub-model performs the reverse operation, that is a projection of $\Re^{\ell}$ to $\Re^{m}$. Thus, in the linear case these two sub-models are characterized by the eigenvectors orthogonal matrix of the data correlation matrix $\widehat{P}$ and the overall model is given by the projection matrix $\widehat{C}^{(\ell)} = \widehat{P}\widehat{P}^{T}$.

In the nonlinear case, the goal is to find two non-linear functions Ψ and Φ. Φ is the nonlinear sub-model of compression to calculate the nonlinear principal components from the original data, and Ψ is the decompression nonlinear sub-model for estimating the original variables from the nonlinear principal components given by the compression sub-model. Therefore, the projection sub-model gives from the data matrix X the principal components T, and the decompression sub-model allows to give the matrix $\widehat{X}$ the estimate of X, based on principal components T. In this case, we can write:

$$T = \Phi(X) \tag{1}$$

Where Φ is a non-linear function equivalent to the eigenvectors matrix *"P"* of the linear PCA. While:
$X \in R^{n\times m}$ the data matrix, and $T \in R^{n\times \ell}$ the principal components matrix. The decompression sub-model provides an estimate $\widehat{x}$ of x from the nonlinear components t (such as x and t are the lines of X and T, respectively):

$$\widehat{X} = \Psi(T) \tag{2}$$

Thus, the data matrix X containing m variables can be expressed as a function of ℓ first nonlinear components.

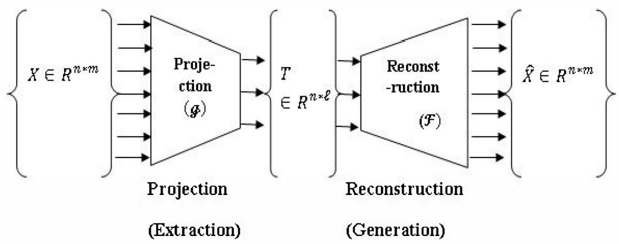


Fig. 2. Basic principles of PCA model

And then the data matrix X can be represented by the estimate $\widehat{X}$ plus the estimation error $\widetilde{X}$ (residual matrix) :

$$X = \widehat{X} + \widetilde{X} = \Psi(T) + E \tag{3}$$

Where $T = \Phi(X)$ is the nonlinear principal components matrix such as $T = [T_1, ...T_\ell]$ and E the residual matrix. The problem is then to identify the nonlinear projection functions Φ and Ψ. Where $\widehat{X} \in R^{n\times m}$ the data reconstruction matrix with $\widehat{X} = \Psi(\mathrm{t})$, Ψ represents the nonlinear function of reconstruction or generation. We note $E\ (k)$ the squared error obtained from the resulting errors on m output neurons, by a learning algorithm of neural network based on the principle of optimization, we search to minimize, by nonlinear optimization methods, the following cost function:

$$\min \sum_{k=1}^{N} \|X(k) - \widehat{X}(k)^{2}\| = \sum_{k=1}^{N} \|X(k) - \Psi(\Phi(\mathrm{X(k)})\|^{2} \tag{4}$$

IV. NEURAL APPROACH OF NLPCA

The nonlinear principal component analysis (NLPCA) based on neural networks, has known a considerable progress and interest in recent years and has been widely used in the field of diagnosis. In this section, we will present a neural network with five layers for extracting nonlinear principal components.

A. Auto-associative neural network

Non linear principal component analysis (NLPCA) which is an extension of linear PCA, has a particular interest in the last years. Most of researchers use a neural approach to calculate the NLPCA model proposed by Kramer [2]. In the case of a single non-linear principal component, the structure of such network is illustrated in figure (Fig. 4). To make the NLPCA, the auto-associative network contains three layers between the input and output variables. A transfer function Ξ1 performs a projection of the input column vector of dimension m to the first hidden layer (coding layer), represented by $h_j^{(x)}$ $(j = 1; :::; r)$ column vector of dimension r (r is the number of neurons in the first hidden layer):

$$h_j^{(x)} = \Xi_1(\sum_{i=1}^{m} v_{ij}^{(x)} + b_j^{(x)}) \tag{5}$$

$V^{(x)}$ is the weight matrix of dimension $(r \times m)$, is a vector containing the r bias parameters. The second transfer function Ξ_2 projects the outputs data of the first hidden layer (coding layer) to the "bottleneck layer" containing a single neuron, which represents the nonlinear principal component t. The transfer function Ξ_1 is generally nonlinear (usually the hyperbolic tangent function or sigmoid function), while the Ξ_2 function is usually the identity function $(\Xi_2\ (x) = x)$:

$$t = \Xi_2(\sum_{i=1}^{r} w_j^{(x)} h_j^{(x)} + \overline{b}^{(x)}) \tag{6}$$

Next, the transfer function Ξ_3, which is a nonlinear function, projects the data from t to the last hidden layer (decoding layer): $h_j^{(t)}$ $(j = 1; : : ; r)$ where r represents the number of neurons in the third hidden layer:

$$h_j^{(t)} = \Xi_3(w_j^{(t)} t + b_j^{(t)}) \tag{7}$$

The last transfer function Ξ_4 is the identity function which projects the outputs data from $h_j^{(t)}$ to $\widehat{x}$: the output column vector of dimension m:

$$\widehat{x}_i = \Xi_4(\sum_{j=1}^{r} v_{ij}^{(t)} h_j^{(t)} + b_i^{(t)}) \tag{8}$$

The cost function $E = ||X(k) - \widehat{X}(k)^2||$ is minimized to find the optimal values of the $V^{(x)}$, $b^{(x)}$, $w^{(x)}$, $\overline{b}^{(x)}$,$w^{(t)}$, $b^{(t)}$, $v^{(t)}$, and $\overline{b}^{(t)}$,

It should be noted that the extraction of principal components can be done in two ways. The first is to extract the principal components sequentially with a single neuron in the middle layer "bottleneck layer" (sequential NLPCA) (Fig. 4). The second is to extract the desired components simultaneously inserting the neurons in the middle layer (NLPCA parallel or simultaneous).

Most learning algorithms of neural networks are the optimization, they seek to minimize, by nonlinear optimization methods, a cost function. This optimization is done iteratively modifying the weights according to the gradient of the cost function (which, in the case of a supervised learning, constitutes a measure of the difference between the real responses of the network and its desired responses). The weights are randomly initialized before the training, and modified iteratively until theobtaininga satisfactory compromise between the accuracy of the approximation on the training set and the accuracy of the approximation on the validation set.

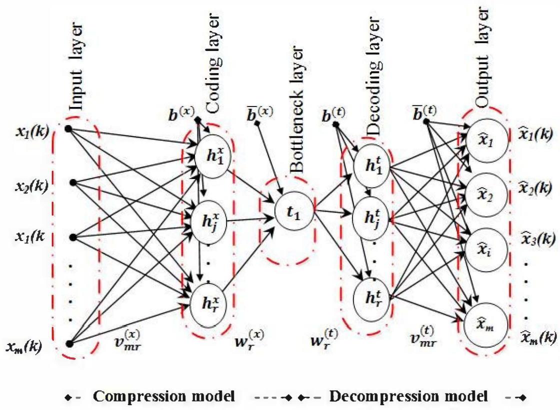


Fig. 3. Auto-associative network with a five layer for extracting one nonlinear principal component proposed by Kramer.

Once the general structure defined, it remains to determine the precise structure of the neural model (network architecture). For this, we must determine the number of necessary hidden layers and the number of neurons in each hidden layer. The number of neurons in the hidden layer is usually determined by performing a cross-validation on validation data set by constructing iteratively the hidden layer, the network growth control is carried out by cross-validation. Note that in this paper, the extraction of principal components is done sequentially by having a single neuron in the middle layer bottleneck layer, where the network is initially trained with a single non-linear principal component. After estimating data from this first non-linear component, we must subtract the obtained result from the starting data set and the extraction operation of the second non-linear component is performed on the obtained residuals . This procedure can be repeated until the desired number of components is reached or the estimation error is below a chosen threshold. So the steps for sequential extraction of nonlinear principal components are:

1. Project the data X in the 1 dimensional space of principal component using a network structure with five layers.
2. Learn the network and estimate $\widehat{X}$ by inverse projection.
3. Calculate the estimation error $E = X - \widehat{X}$. Repeat steps 1 and 2 by taking the matrix E instead of X until a percentage of the variance of X is captured.

V. SENSORS FAULT DETECTION BASED ON SPE INDICATOR

Once the NLPCA model is obtained, we present in this section its use in sensor fault detection. The indicator of detection *SPE* (Squared Prediction Error) performs the fault detection in the residual space. At sample k, is given by:

$$SPE(k) = e(k)e(k)^T \tag{9}$$

$$e(k) = x(k) - \widehat{x}(k) \tag{10}$$

Where, *e (k)* represents the vector of estimation errors. The process is considered abnormal operating (presence a default at the sample *k*) if:

$$SPE(k) > \delta_\alpha^2 \tag{11}$$

Where δ_α^2 is the upper control limit of *SPE(k)*, determined theoretically by BOX [3] , such that $\theta_i = \sum_{j=\ell+1}^{m} \lambda_j^i$, for $i = 1,2,3$ and λ_j is the jth eigenvalue of the matrix Σ to the power of i. The upper control limit theory, for the confidence threshold α given, is then $\delta_\alpha^2 = g\chi^2_{h,\alpha}$:

Where $g = \theta_2/\theta_1$, h=integer (θ_1^2/θ_1) (integer(z) is integer number of z) and $\chi^2_{h,\alpha}$ is the distribution of the χ^2 with h degree of liberty.

VI. LOCALIZATION

After detecting a fault, it is necessary to identify which sensor becomes faulty, it is through the principle of fault localization, for this several methods used in fault localization within of NLPCA. In this paper we present two methods to identify the faults:

A- The contribution plots: 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. In this case the contribution plots $cont_j^{SPE}$ of j^{th} variable at time k is defined by the equation:

$$cont_j^{SPE}(k) = (e_j(k))^2 = (x_j(k) - \widehat{x}_j(k))^2 \tag{12}$$

B- The Non linear data reconstruction principle and Sensor Validity Index (SVI):The principle of reconstruction consists to estimate one of variables denoted x(k) from all other variables, at a given time denoted x_i(k) using the PCA model already obtained, once the estimate of the jthvariable $\widehat{x_j}$ is calculated, we replace the variable $x_i(k)$ by its estimate $\widehat{x_j}$, and we re-estimate this jth variable. this operation is repeated until convergence to the value z_i, using an iterative algorithm, each iteration is an orthogonal projection in a subspace of the principal components, the convergence value z_i is given by:

$$z_i(k) = \xi_i^T \Phi(\Psi(x_i)) \tag{13}$$

where $x_i = [x_1, x_2..\widehat{z_i}..x_m]^T$ and ξ_i is the jth column of the identity matrix.

The *SVI* method is based on the principle of reconstruction. Consists to suspect a faulty sensor and reconstruct the value of the measure based on the PCA model already calculated and the measurements of other sensors. The procedure is repeated for each sensor. The localization is performed by comparison of the detection index before and after reconstruction. The Sensor Validity Index is a measure of sensor performance. It should have a standarized range regardless of the number of principal components, noise, measurement variances or type of faults. The SVI should also distinguish the abnormal operational conditions from the sensor fault situation. The ratio of SPE_j and the SPE can provide these desired properties for the identification of a sensor fault [4]:

$$\eta_j^2 = \frac{SPE}{SPE_j} \tag{14}$$

Where the *SPE* is the global squared prediction error calculated before reconstruction and SPE_j is calculated after the reconstruction of the j^{th}sensor.

VII. APPLICATION

In the field of water, the desired effectiveness in terms of effluent quality and economies of treatment costs has made necessary the modeling, identification and monitoring of biological treatment processes. The Severe requirements on the treatment of nitrogen in wastewater treatment plants made necessary to improve the performance of activated sludge process. Indeed, the strict norms on discharges about treatment plants wastewater are becoming more stringent. In this section and to illustrate the effectiveness of the diagnosis procedure studied, we will present in the section (A) a reduced ASM1(activated sludge model) of WWTP, and in the section (B) a process of a complete station: Annaba WWTP, situated in the northeast of Algeria, in order to apply the diagnosis approaches previously proposed. At this purpose, we demonstrate the validity and applicability of diagnosis procedures using the simulator the reference [11] and a real data collected from WWTP of annaba, the below we present the results of simulation obtained during our studies.

1) Section A: ASM 1 is a model of activated sludge system simulating the phenomena such as the oxidation of carbon, nitrification and denitrification by quantifying the kinetics and stoechiometry of each reaction. On the Base of the ASM1 model, a reduced model proposed by [12] is developed for the activated sludge process in sequential ventilation, the simplified model contains 5 variables, the inflow Q_{in} and incidents concentrations X_{DCO}^{in}, S_{NO}^{in}, S_{NH}^{in} and S_{ND}^{in} present in the influential. The equations which describe the degradation of organic matter in ASM reduced are grouped into equations (15) using the notation proposed by IAWQ1 (HENZE and *al., 1987),* the processing chain is composed of a bioreactor; a clarifier and a sludge recycling loop. The proceed of activated sludge treatment is schematized in Figure fig.5: We Note that the different steps of

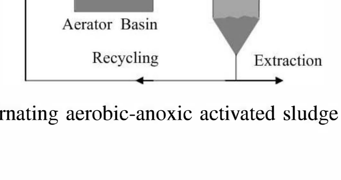


Fig. 4. Alternating aerobic-anoxic activated sludge treatment plant

simplification were made to obtainthe reduced model; the simplifying hypotheses used to reach the reduced model are inspired by Chachuat thesis [12]. The reduced model chosen is described by equations(15):

$$
\begin{aligned}
&\bullet \dot{X}_{DCO} = D^{in}(X_{DCO}^{in} - \frac{K_S}{K_{DCO}} X_{DCO}) - \frac{1}{Y_H}(\rho_1 - \rho_2)\\
&\bullet \dot{S}_{NO} = D^{in}(S_{DCO}^{in} - S_{NO}) - \frac{1-Y_H}{2.86 Y_H}\rho_2 + \frac{1}{Y_A}\rho_3\\
&\bullet \dot{S}_{NH} = D^{in}(S_{NH}^{in} - S_{NH}) - i_{NBH}(\rho_1 + \rho_2) - \frac{1}{Y_A}\rho_3 + \rho_6\\
&\bullet \dot{S}_{ND} = D^{in}(S_{ND}^{in} - S_{ND}) - \rho_6 + \rho_8\\
&\bullet \dot{S}_O = D^{in} S_O - \frac{1-Y_H}{Y_H}{}_1 - 4.57\frac{1}{Y_A}\rho_3 + k_L a(S_O^{sat} - S_O)
\end{aligned} \tag{15}
$$

Where: $\rho_1 = \theta_1 \frac{X_{DC\bullet}}{X_{DC\bullet}+K_{DC\bullet}} \frac{S_\bullet}{S_\bullet + K_{\bullet,H}}$, $\rho_2 = \theta_1 \eta_{NOg} \frac{X_{DC\bullet}}{X_{DC\bullet}+K_{DC\bullet}} \frac{K_{\bullet,H}}{K_{\bullet,H}+S_\bullet} \frac{S_{N\bullet}}{S_{N\bullet}+K_{N\bullet}}$, $\frac{1}{Y_H}\rho_3 = \theta_3 \frac{S_{NH}}{S_{NH}+K_{NH,A}} \frac{S_\bullet}{S_\bullet+K_{\bullet,A}}$, $\rho_4 = b_H X_{B,H}, \rho_5 = b_A X_{B,A}, \rho_6 = \theta_4 S_{ND}$, $\rho_8 = \theta_5 \frac{X_{DC\bullet}}{X_{DCO}+K_{ND}} (\frac{S_\bullet}{S_O+K_{O,H}} + \eta_{NO,h} \frac{K_{\bullet,H}}{K_{\bullet,H}+S_\bullet} \frac{S_{N\bullet}}{S_{N\bullet}+} K_{NO})$.
And;

$\theta_1 = \mu_H X_{B,H}, \theta_2 = (1 - fr_{XI})(\rho_4 + \rho_5), \theta_3 = \frac{\mu_A}{Y_A} X_{BA}$
$\theta_4 = K_a X_{BH}, \theta_5 = K_h \frac{X_{DC\bullet}}{X_S} X_{B,H}, K_{DCO} = K_S \frac{K_{DCO}}{S_S} = \frac{K_S}{f_{SS}}$
$K_{ND} = K_X \frac{K_{DC\bullet}}{X_S} X_{B,H}, D^{in} = \frac{Q_{in}}{V_\bullet}$.

From the reduced model developed in equations (15), we can simulate the evolution of the concentrations of the different components present in the aeration basin of the first biology. Then this simulator, we generate the values, which we will serve as "measures" for implementation of the diagnosis method. In addition, these equations contain a number of terms which must be fixed (parameters relative to the flows, to volumes as well as biochemical parameters intrinsic to the system), the inflow Q_{in} and incidents concentrations X_{DCO}^{in}, S_{NO}^{in}, S_{NH}^{in} and S_{ND}^{in} present in the influential: are equals to the sum of *Sin* and *Cos* of different frequencies and amplitudes.

This reduced model has five state variables X_{DCO}, S_{NO}, S_{NH}, S_{ND} et S_O) (Tab.1) , 16 parameters and Q_{in} and incidents concentrations X_{DCO}^{in}, S_{NO}^{in}, S_{NH}^{in} and S_{ND}^{in}. The parameters sttoichiometric and kinetic are identical to those defined in ASM1 model. However, additional parameters θ_i(i =1;.....; 5), K_{DCO} and K_{ND} are issued from simplifications. The specific values of these parameters are given in Table 2. From a complete simulator of activated sludge secondary treatment of urban effluent, that has been widely used as a reference tool, we take the basin volume V_0 = *1333m* ^{3} the coefficient $K_L a$ relative to the transfer of Oxygen has a value of 240 d^{-1}. *"This simulator was proposed by the working group of the European COST 624"* [11].

Table I: State variables of the reduced model to biological degradation.

Compound	Symbol	Unit
1. Nitrogen as nitrites and nitrates	S_{NO}	gN. m -3
2. Nitrogen as ammonia	S_{NH}	gN. m -3
3. Biodegradable soluble organic nitrogen	S_{ND}	gN. m -3
4.Biodegradable particulate organic nitrogen	X_{ND}	gN. m -3
5. Dissolved oxygen	S_O	$gO_2.m-3$

Table I: Parameters values specific

Parameter	Value
Θ_1	9,956
Θ_2	693
Θ_3	283
Θ_4	124
Θ_5	480
K_{DCO}	220
K_{ND}	258
$X_{B\ A}$	136 g $_{COD.}$ M^{-3}
$X_{B\ H}$	2489 g $_{COD.}$ M^{-3}
X_{ND}	6 g $_{N.}$ M^{-3}

2) Section B: wwtp of Annaba

The wastewater treatment plant of Annaba is the third unit in algeria in terms of its capacity, situated in the municipality of El Bouni, 450 Km from Algiers and 8 Km east of Annaba. The process used is an activated sludge at low load and comprises two processing of traitment, one for water and one for the sludge. In a first phase, the STEP, whose the means flow to purify is 83,620 m^3/d, the commissioning of this infrastructure will enable the reuse of wastewater for irrigation in agriculture (under certain conditions) , industry (ArcelorMittal to 30 000 m3 / d), to the satisfaction of municipal needs (irrigation and street cleaning) and artificial recharge of aquifers located underneath the sea level to avoid the rise of marine fouling, The rate applied to the peak flow leads to up maximal recirculation rate of the $1{,}040 m^3/h$ by file.
We used a real database collected from the WWTP of Annaba. For the monitoring model we selected a measurement vector constructed from (14): x(k) = [S_{O_2inf} S_{O_2eff} MES_{inf} MES_{eff} COD_{inf} COD_{eff} $BOD5_{inf}$ $BOD5_{eff}$ $NH4_{inf}$ $NH4_{eff}$ $NO3_{inf}$ $NO3_{eff}$ $NO2_{inf}$ $NO2_{eff}$]

VIII. Simulation results

In this section, we will give the results obtained from the developed procedure of diagnosis for faults affecting different sensors.
1) Section A:

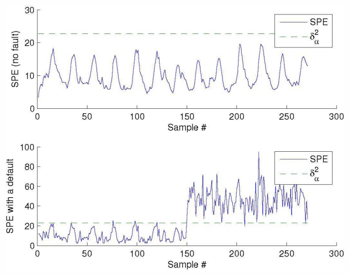


Fig.5 Evolution of the SPE index during normal stat (no fault) and with a default in the variable S_{NH} at sample 150

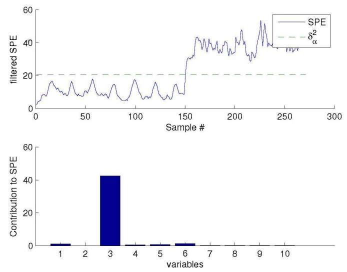


Fig.6 Evolution of filtered SPE and the fault localization affecting the 3^{th} sensor measuring the S_{NH} concentration at sample 150

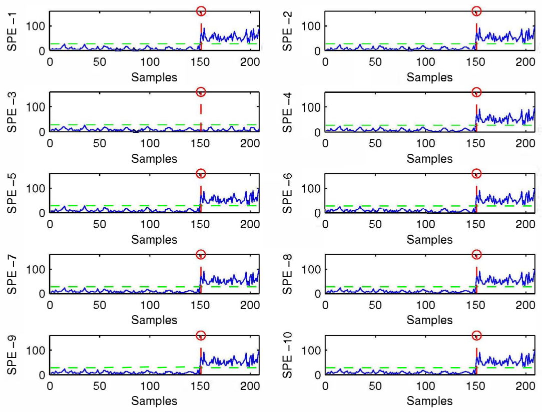


Fig.7 Evolutions of SPE obtained after reconstruction and localization by reconstruction with a fault affecting the 3^{th} sensor measuring the S_{NH}.

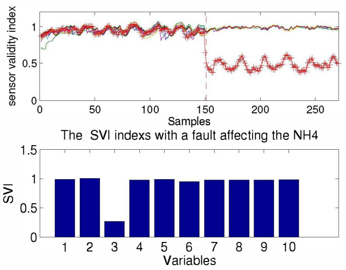


Fig.8 localization from the evolution of SVI with a fault affecting the 3^{th} sensor measuring the S_{NH}

2) Section B:

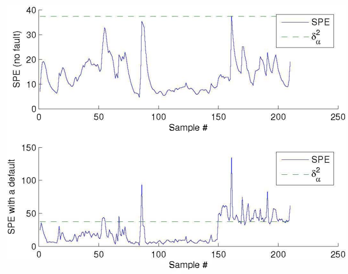


Fig.9 Evolution of the SPE index during normal stat (no fault) and with a default in the variable NH_{4eff} at sample 150

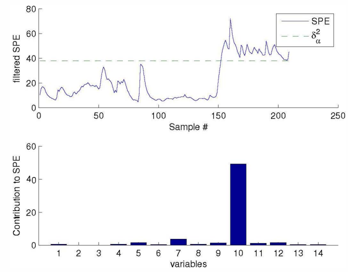
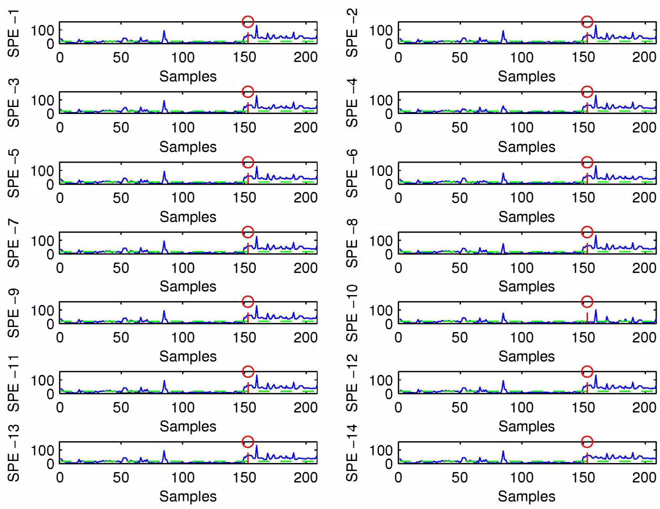
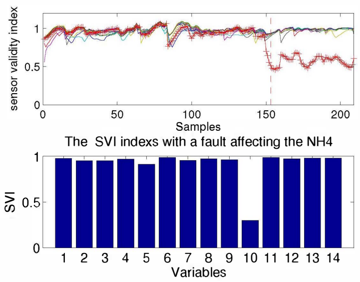


Fig.10 Evolution of filtered SPE and the fault localization affecting the 10^{nd} sensor measuring the NH_4eff concentration at sample 150



Figs.11 Evolutions of SPE after reconstruction and localization by reconstruction with a fault affecting the 10^{th} sensor measuring the NH_{4eff}



Figs.12 localization from the evolution of SVI with a fault affecting the 10^{th} sensor measuring the NH_{4eff}

The figures (Fig. $5_{sectionA}$ and Fig. $9_{sectionB}$) Present the evolution of SPE indicator, that exceeds the upper control limit, it is clear that we detect a fault, at sample 150. For localization anything we can observe in this figures, for this we want to locate the incriminated variable, firstly; with the procedure of contribution plots, secondly; with the method based on reconstruction principle. To avoid false alarms we use the EWMA to filter the effect of outliers and noise, the figures (Fig. $6_{sectionA}$ and Fig. $10_{sectionB}$) present the evolution of the filtered SPE in the presence of a fault affecting the the variables respectively: $S_{NH/Sec(A)}$ and $NH_{4eff/Sec(B)}$, and the fault localization affecting the same variables based on the contribution plots.

According to the figures (Fig. $7_{sectionA}$ and Fig. $11_{sectionB}$) we see that all indicators SPE_i are above their thresholds except statistical, $SPE_{3/sec(A)}$ and $SPE_{10/sec(B)}$ who correspond to the $3^t h$/section (A) and $10^t h$/section(B) variables are lower than its detection limit. It is also noted that the reconstruction of the the $3^t h$/section (A) and $10^t h$/section(B) variables eliminates the defect effect, this expresses that the $3^t h$/section (A) and $10^t h$/section(B) sensor is infected. After the reconstruction the sensor validity index SVI is calculated to locate the fault,the Figures (Fig. $8_{sectionA}$ and Fig. $12_{sectionB}$) present the evolution of SVI with a fault affecting at sample 150 the sensor measuring the variable S_{NH} for section (A) and the variable NH_{4eff} for section(B) these figures we can locate the fault from the decreases variable, which is the 3^{nd} *SVI* corresponding the sensor measuring the S_{NH} concentration for the section (A) and the 10^{th} *SVI* corresponding the sensor measuring the NH_{4eff} concentration for the section (B).

IX. CONCLUSIONS

In recent years, the fault detection and diagnosis methods have been widely developed and used to improve the process operation; particularly the fault detection based on Principal Components Analysis" which does not require prior knowledge about the process mechanism" had known a big progress and has been widely developed. The PCA is a linear modeling tool based on the selection of an optimal number of principal components, its nonlinear extension NLPCA model is obtained using a neural network with five layers in cascade. For monitoring a process, the statistical SPE is used to detect abnormalities, to identify the faulty variables, two diagnosis algorithms are used such as the localization based on the contribution plots and the localization based on the non linear reconstruction principle. The filter applied to the SVI and SPE adds an important feature for sensor fault localization because reduces the effect of false alarms. The principal idea of this article is to apply the diagnosis of sensors operating state used in WTP, the simulation results obtained in this work show the effectiveness of the proposed approaches. Although the efficacy of this method, it can be improved for the better in the future.

References

[1] Hastie T, Principal curves and surfaces. PhD thesis, Stanford University, 1984.

[2] Kramer M. A, Nonlinear principal component analysis using autoassociative neural networks. AIChE Journal, vol. 37, N. 2, pp. 233-243, 1991.

[3] Box G. E. P, Some theorems on quadratique forms applied in the study of analysis of variance problems : Effect of inequality of variance in one-way classification. The Annals of Mathematical Statistics, vol. 25, pp. 290-302, 1954.

[4] Ricardo Dunia et al, Use of principal components analysis for sensor fault identification, Computers chem. Engng, Vol. 20, pp. 713-718, 1996.

[5] George C. Ranger and Frank B. Alt, Choosing principal components for multivariate statistical process control, Communication in Statistics -Theory and Methods . Vol. 25, No. 5, pp.909-922, 1996.

[6] Qin S. J. and Weihua L. Detection, identification and reconstruction of faulty sensors with maximized sensitivity. AIChE Journal, vol. 45, N 9, pp. 1963-1976, 1999.

[7] 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, pp. 69-84, 1985.

[8] Ricardo Dunia et al, Use of principal components analysis for sensor fault identification, Computers chem. Engng, Vol. 20, pp. 713-718, 1996.

[9] T. Lopez-Arenas, A. Pulis and R. Baratti, On-line monitoring of a biological process for wastewater treatment, AMIDIQ, Vol. 3, pp. 51-63, 2004.

[10] Jong-Min Lee, ChangKyoo Yoo, et al. Nonlinear process monitoring using kernel principal component analysis, Chemical Engineering Science, Vol. 59, pp. 223-234, 2004.

[11] J. Alex, J.F. Beteau, J.B. Copp, C. Hellinga, U. Jeppsson, S. Marsili-Libelli, M.N. Pons, H. Spanjers, and H. Vanhooren. Benchmark for evaluating control strategies in wastewater treatment plants. In European Control Conference 1999 , ECC99, Karlsruhe, Germany, August 31- September 3 1999.

[12] B. Chachuat. Mthodologie doptimisation dynamique et de commande optimale des petites stations dpuration boues actives. PhD thesis, Institut National Polytechnique de Lorraine, France, 2001.