

Research Article

Biological Nitrogen Removal Process Monitoring Based on Fuzzy Robust PCA

¹Nabila Heloulou, ¹Messaoud Ramdani and ²Abdenabi Abidi

¹Department of Electronics,

²Department of Engineering, Laboratoire d'Automatique et Signaux de Annaba (LASA),
Université Badji-Mokhtar de Annaba, BP. 12, Annaba 23000, Algeria

Abstract: In this study the Fuzzy Robust Principal Component Analysis (FRPCA) method is used to monitor a biological nitrogen removal process, performances of this method are then compared with classical principal component analysis. The obtained results demonstrate the performances superiority of this robust extension compared with the conventional one. In this method fuzzy variant of PCA uses fuzzy membership and diminish the effect of outliers by assigning small membership values to outliers in order to make it robust. For the purpose of fault detection, the *SPE* index is used. Then the fault localization by contribution plots approach and *SVI* index are exploited.

Keywords: Biological nitrogen removal, fault diagnosis, fuzzy robust PCA, multivariate statistical process control, process monitoring, water treatment plant

INTRODUCTION

During the last few decades, wastewater treatment plant has become essentially a large and complex industrial factory. Increasing requirements on modeling and process monitoring efficiency are becoming more important in order to increase the potential for improved performance and reliability. Several researches concerning the wastewater treatment modeling have been identified in the scientific literature as a useful engineering tool in design, operation and control, where the first technology that was model in this field is the activated sludge process, it has been proved to be a very stable process. Today the Activated Sludge Process (ASP) is the commonly used process for treating municipal as the industrial wastewaters, where the most successful model is the activated sludge process Model No. 1 (Henze *et al.*, 1987). This model triggered the general acceptance of the biological modeling (Tomita *et al.*, 2002) 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, a key element of the WWTP optimization best functioning is the efficient monitoring techniques. Large number of process monitoring including fault detection and diagnosis based on statistical process control has increased, almost exponentially over the last few decades (Yoo *et al.*, 2003; Lee *et al.*, 2004a; Zhao *et al.*, 2004; Lee *et al.*, 2004b; Wang and Romagnoli, 2005; Aguado and Rosen, 2008; Moon *et al.*, 2009;

Corona *et al.*, 2013). Among them the PCA is the most popular method, it has been successfully applied for collected data analysis from systems in the course of operation in order to supervise their behavior. Several extensions of PCA have been investigated and extended in treatment plants. In Haimi *et al.* (2013), authors review the publications focusing on results of PCA application in the biological WWTPs over the last 15 years.

In this study, we investigate the effectiveness of the novel method (FRPCA) compared with the PCA one, for that these two methods will applied on biological process in order to monitor and achieve the earlier fault detection to execute corrective actions before a dangerous occurs in this process. A realistic collection of data is used to validate the technique and conclusions are drawn in the end.

MATHEMATICAL FORMULATION

Traditionnel PCA: Principal Component Analysis (PCA) is a projection-based technique that facilitates a reduction in data dimension through the construction of orthogonal principal components which are linear combinations of the original variables and generally demonstrate the data more feasible in much less dimension (Runger and Alt, 1996). These principal components are ordered so that the first Principal Component variable (PC) in the linear combination has the greatest variance, the second PC is the linear combination with the next greatest variance. The

Corresponding Author: Nabila Heloulou, Department of Electronics, Laboratoire d'Automatique et Signaux de Annaba (LASA), Université Badji-Mokhtar de Annaba, BP. 12, Annaba 23000, Algeria

This work is licensed under a Creative Commons Attribution 4.0 International License (URL: <http://creativecommons.org/licenses/by/4.0/>).

remaining PCs are defined similarly. It is often the case that a small number of principal components is sufficient to account for most of the structure in the data. The transformation is defined by:

$$X = \sum_{j=1}^{\ell < m} t_j p_j^T + E \tag{1}$$

where, p_j is a $m \times 1$ component (loading vectors), they are defined here as being orthonormal and so they become the eigenvectors of the data covariance matrix $X^T X$. t_j is the $n \times 1$ score vector corresponding to j^{th} variable. m is the number of principal components retained and E is the residual error. A new sample vector X can be decomposed into two parts:

$$X = TP^T + \tilde{X} \tag{2}$$

where, $\tilde{X} = \tilde{T}\tilde{P}^T$ is the residual matrix. For more information, the reader is referred to Jackson (1991), Jolliffe (1986) and Wold *et al.* (1987).

Fuzzy robust principal component analysis: Fuzzy Robust Principal Component Analysis algorithms are designed to solve two majors problems faced by the PCA approach, first the problem of the real applications where the data come in the on line way, the second one is the sensitivity to outliers (Xu and Yuille, 1995). The robust rules developed in Xu and Yuille (1995) make FRPCA techniques essential when better accuracies are yielded when the algorithms are used. The FRPCA algorithms used here were introduced in Yang and Wang (1999), the derived nonlinear case was proposed in Luukka (2009). The algorithms proposed by Yang and Wang (1999) in their paper are based on Xu and Yuille algorithms (Xu and Yuille, 1995). The ability of these algorithms lay in the way they deal with outliers removal in a given data (Luukka, 2009). Yang and Wang (1999) defined a fuzzy objective function which includes Xu and Yuille's as crisp special cases. We present a brief description of the theory by Xu and Yuille, also the modification proposed by Yang and Wang. Xu and Yuille (Yang and Wang, 1999) proposed an optimization function with an energy measure $e(x_i)$ subject to the membership set $u_i \in \{0, 1\}$ given as:

$$E(U, w) = \sum_{i=1}^n u_i e(x_i) + \eta \sum_{i=1}^n (1 - u_i) \tag{3}$$

where, $X = \{x_1, x_2, \dots, x_n\}$ is the data set, $U = \{u_i | i = 1, \dots, n\}$ is the membership set and η is the threshold. The variable u_i serves to decide whether x_i is an outlier or a sample. When $u_i = 1$, the portion of energy contributed by the sample x_i is taken into consideration; otherwise x_i

is considered as an outlier (Moon *et al.*, 2009). The goal is to minimize $E(U, w)$ with respect to u_i and w . Since u_i is the binary variable and, w is the continuous variable, the optimization with gradient descent approach is hard to solve using gradient descent. To overcome the problem, the minimization problem was transformed to maximization of Gibbs distribution with the use of a partition function. The new problem thus looks like below:

$$P(U, w) = \frac{\exp(-vE(U, w))}{Z} \tag{4}$$

where Z is the partition function ensuring $\sum_U \int_w P(U, w) = 1$. The measure could be one of the following functions:

$$e_1(x_i) = \|x_i - w^T x_i w\|^2, \tag{5}$$

$$e_2(x_i) = \|x_i\|^2 - \frac{\|w^T x_i\|^2}{\|w\|^2} = x_i^T x_i - \frac{w^T x_i x_i^T w}{w^T w} \tag{6}$$

The gradient descent rules for minimizing $E_1 = \sum_{i=1}^n e_1(x_i)$ and $E_2 = \sum_{i=1}^n e_2(x_i)$ are:

$$w^{new} = w^{old} + \alpha_t [y(x_i - u) + (y - v)x_i], \tag{7}$$

$$w^{new} = w^{old} + \alpha_t \left(x_i y - \frac{w}{w^T w} y^2 \right). \tag{8}$$

where α_t is the learning rate, $y = w^T x_i$, $u = yw$ and $v = w^T u$. Oja presented a nonlinear PCA (Oja, 1995), where,

$$e_3(x_i) = \|x_i - w^T g(y)\| \tag{9}$$

and where $y = x_i^w$ and g can be chosen as nonlinear function. In this case the weight updating would be:

$$w^{new} = w^{old} + \alpha_t (x_i e^T w^{old} F + e_3(x_i) g(y)) \tag{10}$$

where, $F = \frac{d}{dy}(g(y))$

The fuzzy variant of the objective function Eq. (1) is proposed by Yang and Wang, they adopt fuzzy memberships by altering the membership set u_i , with a factor called the fuzziness variable denoted by m in the next equation below, thus the objective function was stated as:

$$E = \sum_{i=1}^n u_i^m e(x_i) + \eta \sum_{i=1}^n (1-u_i)^m \quad (11)$$

Subject to $u_i \in [0, 1]$ and $m \in [0, 1]$. Now u_i being the membership of x_i belonging to data cluster and $(1-u_i)$ is the membership of x_i belonging to noise cluster. m is the so called fuzziness variable. In this case, $e(x_i)$ measures the error between x_i and the class center. This idea is similar to the C-means algorithm (Oja, 1982). Since u_i is now a continuous variable the difficulty of a mixture of discrete and continuous optimization can be avoided and the gradient descent approach can be used. The derivatives of (7) with respect to both u_i and w were found:

$$\frac{\delta E}{\delta u_i} = \frac{\delta}{\delta u_i} \left(\sum_{i=1}^n u_i^m e(x_i) + \eta \sum_{i=1}^n (1-u_i)^m \right) \quad (12)$$

Setting the derivative to zero gives the solution as:

$$u_i = \frac{1}{1 + \left(\frac{e(x_i)}{\eta} \right)^{1/(m-1)}} \quad (13)$$

Using this result in the objective function and simplifying, we obtain:

$$E = \sum_{i=1}^n \left(\frac{1}{1 + \left(\frac{e(x_i)}{\eta} \right)^{1/(m-1)}} \right)^{m-1} e(x_i) \quad (14)$$

The gradient with respect to w is:

$$\frac{\delta E}{\delta w} = \beta(x_i) \left(\frac{\delta e(x_i)}{\delta w} \right) \quad (15)$$

where,

$$\beta(x_i) = \left(\frac{1}{1 + \left(\frac{e(x_i)}{\eta} \right)^{1/(m-1)}} \right)^m \quad (16)$$

And m is the fuzziness variable. If $m = 1$ the fuzzy membership reduces to the hard membership and can be determined by following rule:

$$u_i = \begin{cases} 1 & \text{if } (e(x_i) < \eta) \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

Now η is a hard threshold in this situation. There is no general rule for the setting of m , but most papers set $m = 2$. In Yang and Wang (1999), authors derived the three following algorithm for the optimization procedure:

FRPCA1 algorithm:

- Step 1:** Initially set the iteration count $t = 1$, iteration bound T , learning coefficient $\alpha_0 \in (0, 1]$ soft threshold η to a small positive value and randomly initialize the weight w .
- Step 2:** While t is less than T , perform the next steps 3 to 9.
- Step 3:** Compute $\alpha_t = \alpha_0 (1-t/T)$ set $i = 1$ and $\sigma = 0$.
- Step 4:** While $i = 1$ is less than T , do steps 5-8.
- Step 5:** Compute $y = w^T x_i$, $u = yw$ and $v = w^T u$.
- Step 6:** Update the weight:

$$w^{new} = w^{old} + \alpha_t \beta(x_i) [y(x_i - u) + (y - v)x_i]$$

- Step 7:** Update the temporary count $\sigma = \sigma + e_1(x_i)$.
- Step 8:** Add 1 to i .
- Step 9:** Compute $\eta = \left(\frac{\sigma}{n} \right)$ and add 1 to t .

FRPCA2 algorithm: The same as FRPCA1 except steps 6-7:

- Step 6:** Update the weight:

$$w^{new} = w^{old} + \alpha_t \beta(x_i) \left(x_i y - \frac{w}{w^T w} y^2 \right)$$

- Step 7:** Update the temporary count: $\sigma = \sigma + e_2(x_i)$

FRPCA3 algorithm: Follow the same as FRPCA1 except steps 6-7:

- Step 6:** Update the weight:

$$w^{new} = w^{old} + \alpha_t \beta(x_i) (x_i y - w y^2)$$

- Step 7:** Update the temporary count: $\sigma = \sigma + e_i(x_i)$

As we can remark, the three algorithms are slightly different but we have applied the new Nonlinear FRPCA3 proposed in Luukka (2009), where the change is in the way of updating the weight and count.

New nonlinear FRPCA3 algorithm: The same as FRPCA3 except steps 6-7:

Step 6: Compute $g(y)$, $F = \frac{d}{dy}(g(y))$, $e_3(x_i) = x_i - w^{old} g(y)$.

Update the weight:

$$w^{new} = w^{old} + \alpha_i \beta(x_i) (x_i e_3(x_i)^T w^{old} F + e_3(x_i) g(y))$$

Step 7: Update the temporary count: $\sigma = \sigma + e_3(x_i)$

The weight w in the updating rules converges to the principal component vector almost surely (Oja, 1982; Oja and Karhunen, 1985).

FAULT DETECTION AND IDENTIFICATION

Fault detection: Once a FRPCA model is built, it is necessary to have a criterion to judge whether this model is valid for control. The multivariate statistical process monitoring makes use of these criteria. The index, which is called the Squared Prediction Error (SPE), also known as Q , is used here, for data tests FRPCA based model. The SPE is given by:

$$Q = SPE(k) = \sum_{j=1}^{\ell} (x_j(k) - \hat{x}_j(k))^2 \quad (18)$$

where, $x_j(k)$ is an FRPCA input and $\hat{x}_j(k)$ is the prediction of $x_j(k)$ from the FRPCA model. With a control limit δ^2 as:

$$\delta_{\alpha}^2 = g \chi_{h,\alpha}^2 \quad (19)$$

where, α is the confidence limit and:

$$g = \frac{\theta_2}{\theta_1} \quad (20)$$

$$h = \frac{\theta_1^2}{\theta_2} \quad (21)$$

where, $\theta_1 = \sum_{i=\ell+1}^n \lambda_i$, $\theta_2 = \sum_{i=\ell+1}^n \lambda_i^2$ and λ_i is the i^{th} eigenvalue of the covariance matrix. The control limit is calculated from reference data. If the SPE is above its control limit, the system is considered in a faulty state, this is explained by the change of correlation structure of the process variables.

Fault identification: When a fault is detected, it is necessary to identify the variable which is at issue; this is name the fault localization.

Several methods for located the variable in defects were proposed in literature, in this study we are

interesting by the techniques of contribution plots and reconstruction fault.

Contribution plots: The classical approach used in fault identification by PCA is based on the calculated contributions to the index of detection (Alcala and Qin, 2009), so the variable with a largest contribution to the detection indicator is the variable incriminating in the cause of detection index SPE. A contribution of the j^{th} variable to the SPE-statistic is defined as:

$$C_j^{SPE} = (x_j - \hat{x}_j)^2 \quad (22)$$

Reconstruction fault: The principle of reconstruction consists in estimating one of the variables of data vector $x(k)$ at a given time denoted $x_i(k)$, using the PCA model and others variables. There are three different approaches that can be used for reconstructing of faulty sensor, in this study the iterative reconstruction algorithm was used to reconstruct faulty sensor. Assumed that fault direction is known, the reconstruction \hat{x}_i of the i^{th} variable from the selected number of principal component by iterative technique whereby the value of the faulty sensor is replaced by the predicted value is given by:

$$\hat{x}_i = \begin{bmatrix} c_{-i}^T & 0 & c_{+i}^T \end{bmatrix} x + c_{ii}^T \hat{x}_i^{old} \quad (23)$$

where, \hat{x}_i^{old} can be considered as a projection of x on the PCs, it can be calculated as by: $\hat{x} = Cx$, where $C = PP^T$ and $\begin{bmatrix} c_{-i}^T & 0 & c_{+i}^T \end{bmatrix}$ is a vector of matrix C which the i^{th} column of c_{ii} equation: is replaced by 0. The iterative process converges to the following formula:

$$x_i^* = \frac{\begin{bmatrix} c_{-i}^T & 0 & c_{+i}^T \end{bmatrix}}{1 - c_{ii}} \quad (24)$$

where, $c_{ii} \neq 1$, In the case of $c_{ii} = 1$, the i^{th} cannot be reconstructed by this method.

In the case of faulty sensor, we have a significant reduction in SPE before and after reconstruction. However, in some situations, reducing the SPE may affect all entries, which makes the faulty sensor unidentifiable. A Sensor Validity Index (SVI) was introduced. Assuming that only one sensor fault occur in the system process (Dunia et al., 1996), this index determine the status of each sensor. It can be defined as:

$$\eta_i^2 = \frac{SPE(x_i^*)}{SPE(x)} \quad (25)$$

where, x_i^* is a reconstructed vector.

Apparently, η_i^2 ranges between (0, 1), because the $SPE(x_i) \geq SPE(x)$. When η_i^2 is close to 0, it indicates that the i^{th} sensor is faulty. On the other hand, when η_i^2 is close to 1, it means that the sensor variations is consistent with others ones.

SIMULATED ACTIVATED SLUDGE MODEL

In this section, an Activated Sludge Process (ASP) for nitrogen removal is presented. This process is well described in Lopez-Arenas *et al.* (2004). The basic design of an ASP is shown in Fig. 1. In its basic configuration, the activated sludge process consists of two reactors and a settler. To enhance the nitrogen removal, anoxic/anaerobic processes operate alternately. During the aerobic phase ammonia nitrogen is converted into nitrite by Nitrosomas and subsequently into nitrate by Nitrobacters, in the anoxic phase the produced nitrate is converted into harmless nitrogen gas. As illustrated in Fig. 1, before it enters the aeration reactor, raw wastewater Q_{in} is passed by the anoxic zone, afterward the influent flow Q_{out} is fed into a settler to separate the stream into the clean water and sludge, the major part of it is recycled to reactor Q_r and a small part is wasted Q_w . The actual process model is based on the Activated Model Sludge No. 1 (ASM1) by Kim *et al.* (2011). It was adopted with two modifications (Lopez-Arenas *et al.*, 2004):

- The nitrification is modeled by a two step process (the conversion of nitrite to nitrate by the nitrosoma bacteria and the conversion of nitrite to nitrate by the nitrobacters).
- The hydrolysis of rapidly biodegradable substrate is included.

Then the resulting biodegradation model consists of 18 state variables (particles and soluble concentrations) and 30 model parameters. However it is

possible to reduce the model, such model is proposed by Gomez-Quintero *et al.* (2000).

This model consists of 8 states variables: dissolved oxygen ($S_{O_2}^p, S_{O_2}^n$), nitrate ($S_{NO_3}^p, S_{NO_3}^n$), ammonia ($S_{NH_4}^p, S_{NH_4}^n$) and biodegradable substrate concentrations (S_S^p, S_S^n), for each reactor zone (p and n denote pre-denitrification and nitrification respectively).

Rewritten the reduced model in state format is given as:

$$\begin{aligned} \dot{x}_1 &= d_1 d_4 + d_5 x_5 - (d_4 + d_5) x_1 - A x_4 r_1^p r_2^p + \alpha_2 r_3^p r_4^p \\ &:= f_1(x_1, x_2, x_3, x_4, x_5, d, p) \\ \dot{x}_2 &= d_5 x_6 - (d_4 + d_5) x_2 + k_{La}^p (S^{OST} - x_2) - E x_4 r_5^p - F r_3^p r_4^p \\ &:= f_2(x_2, x_3, x_4, x_6, d, p) \\ \dot{x}_3 &= d_2 d_4 + d_5 x_7 - (d_4 + d_5) x_3 - B x_4 (r_5^p + r_1^p r_2^p) - \alpha_2 r_3^p r_4^p + \alpha_3 \\ &:= f_3(x_1, x_2, x_3, x_4, x_7, d, p) \\ \dot{x}_4 &= d_3 d_4 + d_5 x_8 - (d_4 + d_5) x_4 + (\alpha_4 - D x_4) r_5^p + (C - D x_4) r_1^p r_2^p \\ &:= f_4(x_1, x_2, x_4, x_8, d, p) \\ \dot{x}_5 &= (d_4 + d_5)(x_1 - x_5) - A x_8 r_1^n r_2^n + \alpha_2 r_3^n r_4^n \\ &:= f_5(x_1, x_5, x_6, x_7, x_8, d, p) \\ \dot{x}_6 &= (d_4 + d_5)(x_2 - x_6) + k_{La}^n (S^{OST} - x_6) - E x_8 r_5^n - F r_3^n r_4^n \\ &:= f_6(x_2, x_6, x_7, x_8, d, p) \\ \dot{x}_7 &= (d_4 + d_5)(x_3 - x_7) - B x_8 (r_5^n + r_1^n r_2^n) - \alpha_2 r_3^n r_4^n + \alpha_3 \\ &:= f_7(x_3, x_5, x_6, x_7, x_8, d, p) \\ \dot{x}_8 &= (d_4 + d_5)(x_4 - x_8) + (\alpha_4 - D x_8) r_5^n + (C - D x_8) r_1^n r_2^n \\ &:= f_8(x_4, x_5, x_6, x_8, d, p) \end{aligned}$$

So, this model consists of eight state variables:

$$x = [x_1, \dots, x_8]^T = [S_{NO_3}^p, S_{O_2}^p, S_{NH_4}^p, S_S^p, S_{NO_3}^n, S_{O_2}^n, S_{NH_4}^n, S_S^n]^T$$

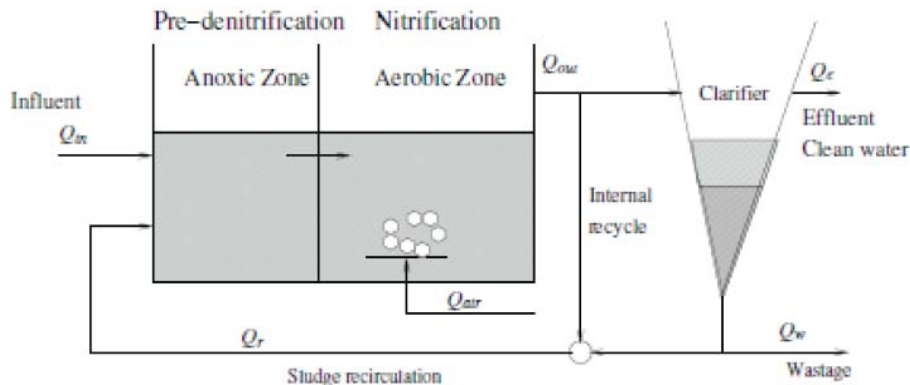


Fig. 1: Schematic of a typical wastewater treatment plant

where,

$$A = \alpha_1 (1 - Y_H), B = \alpha_1 i_{XB}, C = \alpha_4 \eta_H$$

$$D = \alpha_1 / Y_H, E = \alpha_1 (1 - Y_H) / Y_H, F = 4.57 \alpha_2$$

And six exogenous inputs represent the influents concentrations and flow rates:

$$d = [d_1, \dots, d_6]^T = [S_{NO_3}^{in}, S_{NH_4}^{in}, S_S^{in}, Q_{in} / V, Q_r / V, Q_w / V]^T$$

The model needs of only five reaction rates given by:

$$r_1^{p/n} = \frac{S_{NO_3}^{p/n}}{K_{NO_3} + S_{NO_3}^{p/n}}, r_2^{p/n} = \frac{K_{O_2,H}}{K_{O_2,H} + S_{O_2}^{p/n}}$$

$$r_3^{p/n} = \frac{S_{NH_4}^{p/n}}{K_{NH_4} + S_{NH_4}^{p/n}}, r_4^{p/n} = \frac{S_{O_2}^{p/n}}{K_{O_2,A} + S_{O_2}^{p/n}},$$

$$r_5^{p/n} = \frac{S_{O_2}^{p/n}}{K_{O_2,H} + S_{O_2}^{p/n}}$$

And it has twelve model parameters:

$$p = [p_1, \dots, p_{12}] = [Y_H, i_{XB}, K_{O_2,H}, K_{O_2,A}, K_{NO_3}, K_{NH_4}, \eta_g, \eta_H, \alpha_1, \alpha_2, \alpha_3, \alpha_4]^T$$

RESULTS AND DISCUSSION

Simulated activated sludge process: In our case study, a formulated process monitoring system to the problem of faulty sensor consists on estimating the measurement vector Y given as:

$$Y = [y_1, \dots, y_7]^T = [S_{NO_3}^p, S_{O_2}^p, S_{NO_3}^n, S_{O_2}^n, u_1, u_2, u_3]^T$$

where, $u_1 = Q_{in}$, $u_2 = Q_r$, $u_3 = q_{air}$: aeration rate which affect the process through the oxygen mass transfer coefficient k_{La} .

Now, the algorithms for process monitoring, including fault detection and identification are applying to the process. The *SPE* for data in faulty state is given in Fig. 2 to 5 indicates clearly that the fault sensor is $sx = 3$ since it's the sensor validity index drops below a certain threshold δ .

Case study:

The WWTP of Annaba city: In this case study, the wastewater treatment plant of Annaba city (situated in the North-East of Algeria) has two principal stages.

The primary stage: Which includes bar racks, grit chamber, de-oiling and sand filters whose objective is the removal of solids.

The secondary stage: Whose objective is the biological treatment of the organic load and which is essential to remove the organic matter present in the incoming wastewater. The organic matter serve as food for the micro-organisms culture as it grows. Biological treatment is classified by the type of micro-organisms that are used in the removal of the organic pollutants and is either aerobic, anaerobic or both.

Interest is normally focused on secondary biological treatment and the WWTP considered in this study involves extended aeration and it is well

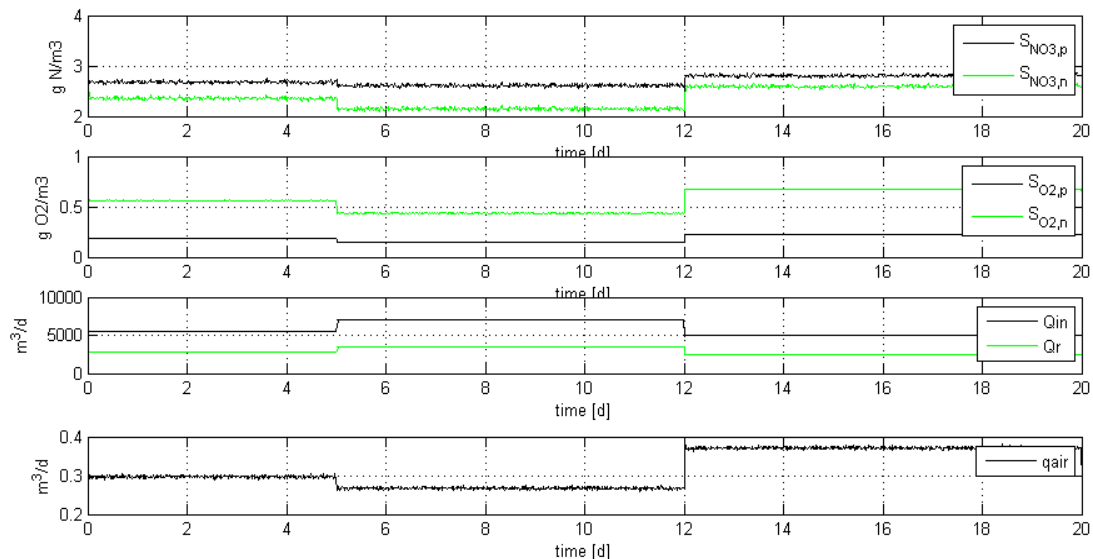


Fig. 2: Time evolution of the selected variables

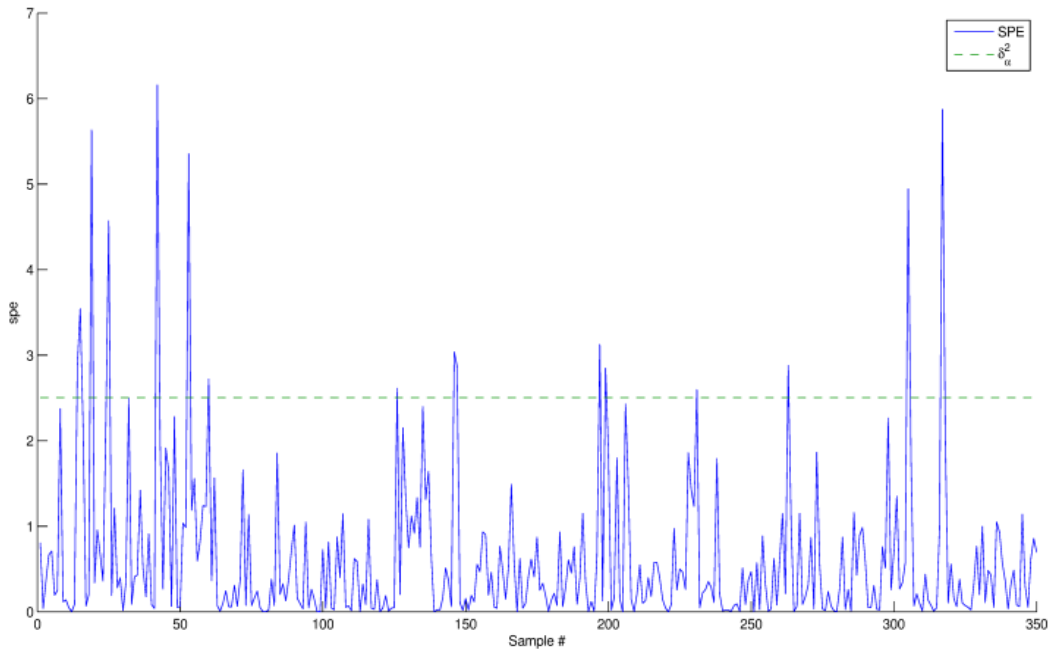


Fig. 3: SPE for normal data

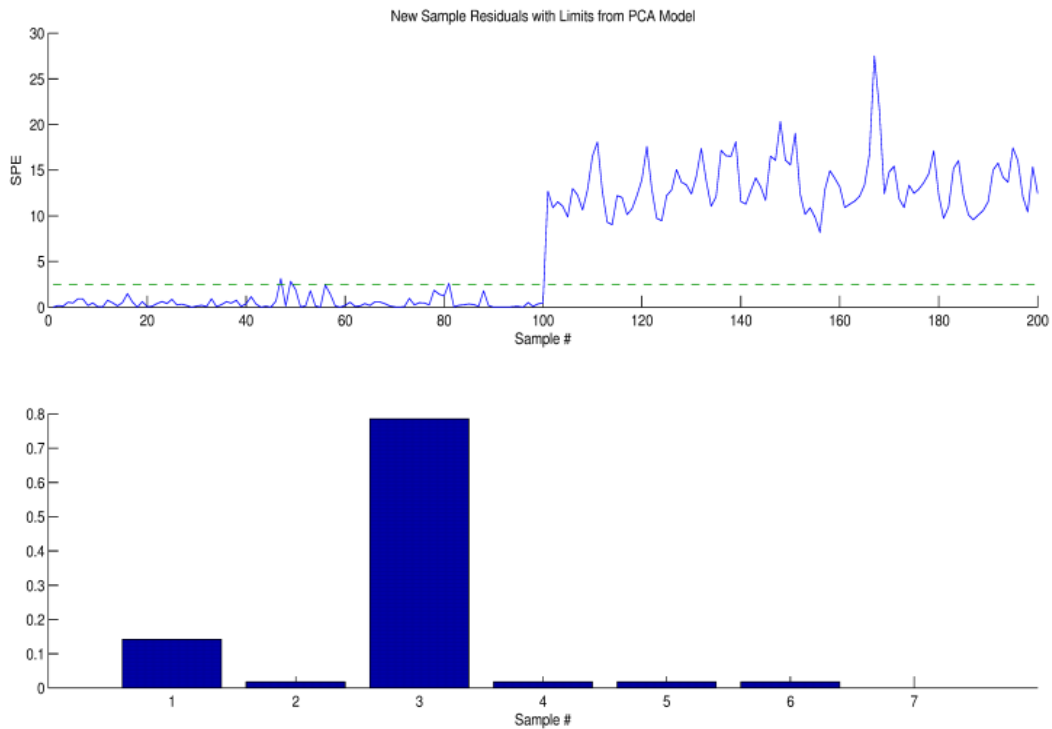


Fig. 4: SPE for data in faulty state, fault isolation using contribution to SPE (fault in the first sensor, $S_{NO_3}^p$)

instrumented. A schematic of the plant is shown in Fig. 6. The actual data is collected in a period of three months of the year 2011. A data matrix X has been formed of $N = 180$ observations which represent a normal operation of the process. 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. Figure 7 represents the statistical SPE of the data set during normal operation process to a confidence level for 95%.

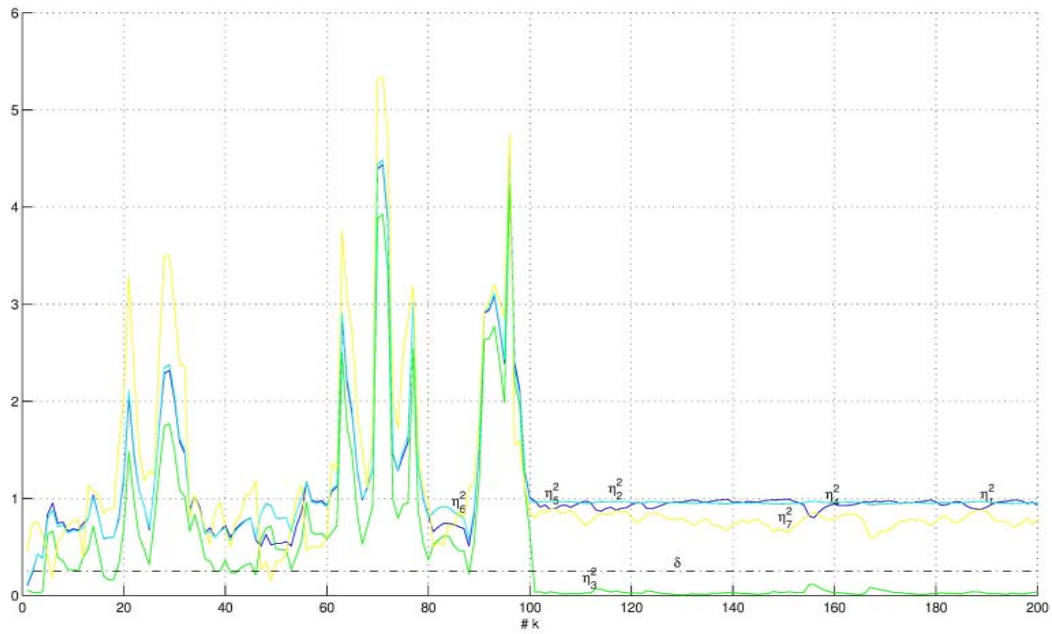


Fig. 5: SVI for the FRPCA with a fault in sensor $s_x = 3$

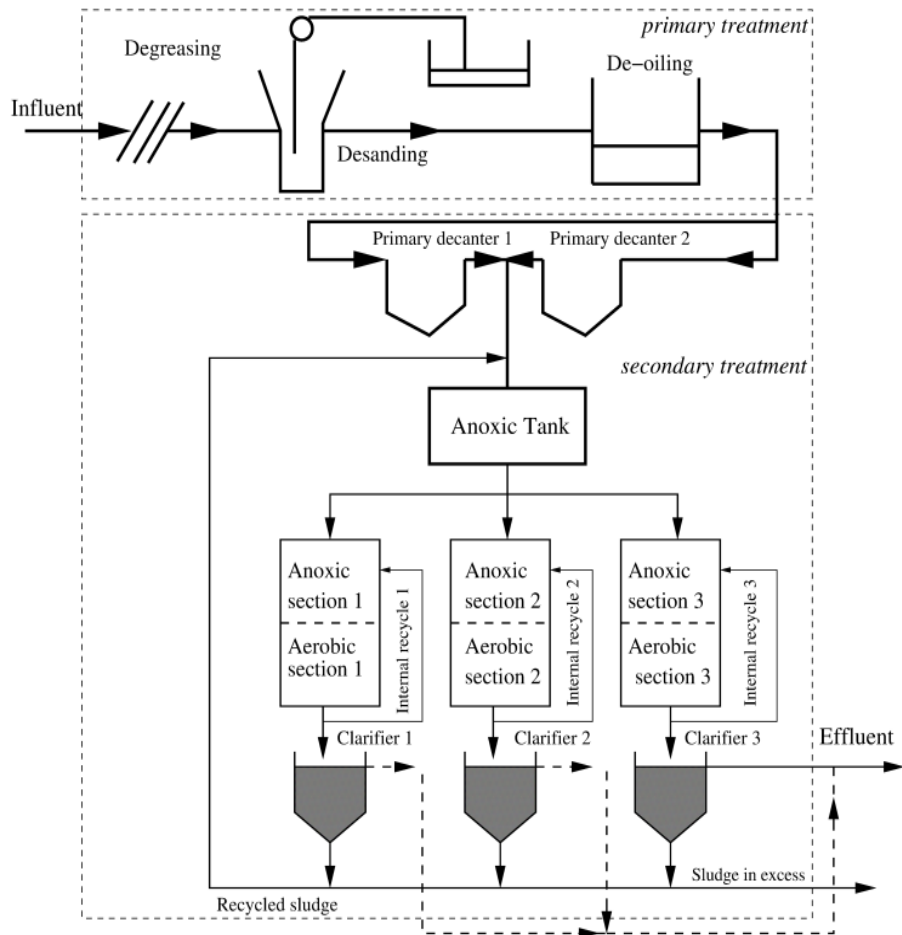


Fig. 6: Schematic of the wastewater treatment plant of Annaba city

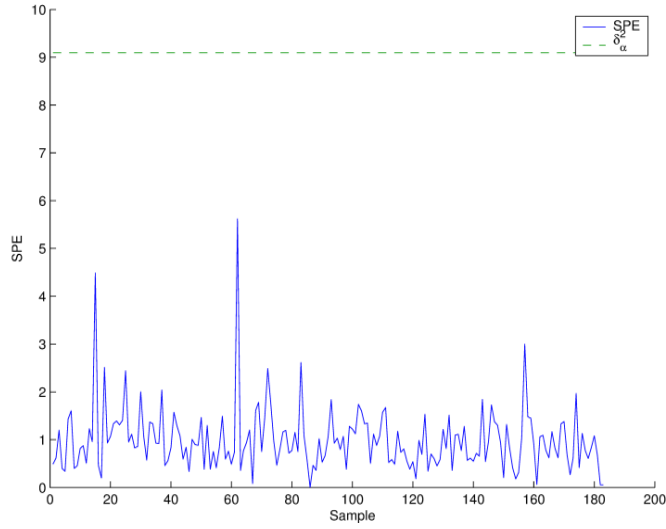


Fig. 7: Statistical SPE to $\alpha = 0.05$

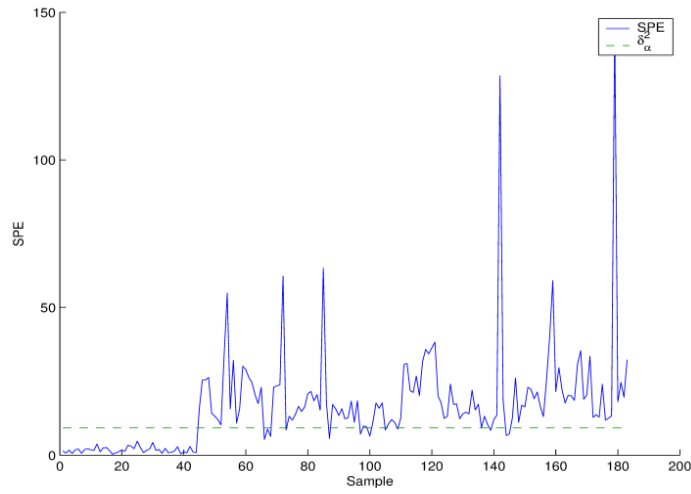


Fig. 8: Statistical SPE to $\alpha = 0.05$

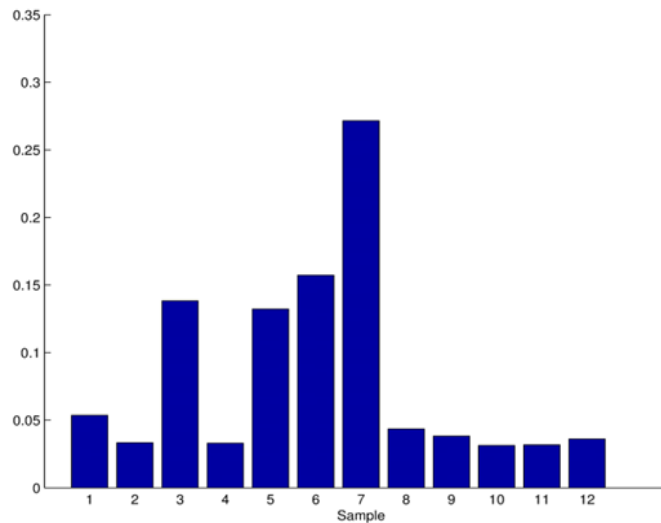


Fig. 9: Localization of defect by the contributions to the SPE index

After the construction of multivariate statistical monitoring model from normal data and in order to illustrate the fault diagnosis performance, an offset fault affecting the x_7 , ($x_7 \equiv S_{O_2,eff}$ oxygen at the output of WWTP) is simulated, from the sample $k = 45$ as a 10% of the maximum amplitude of variation of the variable x_7 . Figure 8 shows the statistical SPE . We note that the statistical SPE indicated in Fig. 9 immediately allows the detection of the fault. To identify the faulty sensor, it was exploited the contributions approach to detection index SPE .

CONCLUSION

In many industries, it is important to determine when significant adverse process changes occur. The idea is to discover these changes while they are still relatively minor, before substandard product or significant pollution is produced. This study discusses the use of robust process control method for the purpose of monitoring wastewater data. It has been previously shown that traditional statistical process control methods are sensitive to noise. The present study investigates the use of robust multivariate statistical process monitoring for biological nutrient plant. It has been shown that the FRPCA is efficient multivariate statistical process monitoring tool. Although the good performance of this method, it can be improved for the better in the future, especially to cover a wide range of normal operating conditions due to changing influent water parameters and organic load.

NOMENCLATURE

d	: Exogenous input
E	: Residual matrices
f	: Model map
i_{XB}	: Mass N/mass COD in biomass
k_{La}	: Oxygen mass transfer coefficient (d^{-1})
$K_{O_2,H}$: Aerobic oxygen half-saturation coefficient
$K_{O_2,A}$: Aerobic/anoxic oxygen half-sat coefficient
K_{NO_3}	: Nitrate half-saturation coefficient
K_{NH_4}	: Ammonia half-saturation coefficient
m	: Number of retained principal components
P	: Loading matrix
p_j	: Loading vector
p	: Model parameter
Q	: Q statistic
r_i	: i -th reaction rate ($1 \leq i \leq 5$)
S_{NO_3}	: Nitrate concentration ($g\ N/m^3$)
S_{O_2}	: Dissolved oxygen concentration ($g\ O_2/m^3$)
S_{NH_4}	: Ammonia concentration ($g\ N/m^3$)
S_S	: Biodegradable substrate conc. ($g\ COD/m^3$)
S_{OST}	: Dissolved oxygen saturation conc. ($g\ O_2/m^3$)
SVI	: Sensor validity index

SPE	: Squared prediction error
T	: Score matrix
t_j	: Score vector
u_i	: Membership set
V	: Reactor volume (m^3)
$X \in \mathbb{R}^{n \times m}$: Matrix of data representing the normal functioning of the system
x	: Process state
x_i^*	: Reconstructed value of measurement x_i
Y_H	: Heterotrophic yield
y	: Measured output

Greek symbols:

α_i	: i -th reduced model parameter ($1 \leq i \leq 4$)
η_g	: Correction factor for anoxic growth
η_H	: Correction factor for anoxic hydrolysis
δ^2	: Confidence limit for Q statistic

Subscripts:

in	: Influent
r	: Return activated sludge, RAS
w	: Waste activated sludge, WAW
out	: Reaction exit
eff	: (Clean effluent)

Superscripts:

P	: Pre-denitrification
n	: Nitrification
^	: Estimated

ACKNOWLEDGMENT

The work reported in this study was supported by the Minister for higher education and research, within the research project under the grant J0201120100088 related to the advanced control and monitoring of biological wastewater treatment plants.

REFERENCES

- Aguado, D. and C. Rosen, 2008. Multivariate statistical monitoring of continuous wastewater treatment plants. Eng. Appl. Artif. Intel., 21(7): 1080-1091.
- Alcala, C.F. and S.J. Qin, 2009. Reconstruction-based contribution for process monitoring. Automatica, 45(7): 1593-1600.
- Corona, F., M. Mulas, H. Haimi, L. Sundell, M. Heinonen and R. Vahala, 2013. Monitoring nitrate concentrations in the denitrifying post-filtration unit of a municipal wastewater treatment plant. J. Process Contr., 23(2): 158-170.
- Dunia, R., S.J. Qin, T.F. Edgar and T.J. McAvoy, 1996. Use of principal components Analysis for sensor fault identification. Comput. Chem. Eng., 20: 713-718.

- Gomez-Quintero, C., I. Queinnec and J.P. Babary, 2000. A reduced nonlinear model of an activated sludge process. *Proceeding of International Symposium on Advanced Control of Chemical Processes*, 2: 1037-1402.
- Haimi, H., M. Mulas, F. Corona and R. Vahala, 2013. Data-derived soft-sensors for biological wastewater treatment plants: An overview. *Environ. Model. Software*, 47: 88-107.
- Henze, M., C.P. Leslie Grady, W. Gujer, G.V.R. Maris and T. Matsuo, 1987. Activated sludge process model No. 1. Technical Report IAWQ Scientific and Technical Report 1.
- Jackson, J.E., 1991. *A User's Guide to Principal Components*. John Wiley and Sons, Inc., New York.
- Jolliffe, L.T., 1986. *Principal Component Analysis*. Springer Verlag, New York.
- Kim, Y.M., H.U. Cho, D.S. Lee, D. Park and J.M. Park, 2011. Influence of operational parameters on nitrogen removal efficiency and microbial communities in a fullscale activated sludge process. *Water Res.*, 45: 5785-5795.
- Lee, J.M., C. Yoo and I.B. Lee, 2004b. Statistical monitoring of dynamic processes based on dynamic independent component analysis. *Chem. Eng. Sci.*, 59: 2995-3006.
- Lee, J.M., C. Yoo, S.W. Choi, P.A. Vanrolleghem and I.B. Lee, 2004a. Nonlinear process monitoring using kernel principal component analysis. *Chem. Eng. Sci.*, 59: 223-234.
- Lopez-Arenas, T., A. Pulis and R. Baratti, 2004. On-line monitoring of a biological for wastewater treatment. *Rev. Mex. Ing. Quim.*, 3: 51-63.
- Luukka, P., 2009. Classification based on fuzzy robust PCA algorithms and similarity classifier. *Expert Syst. Appl.*, 36: 7463-7468.
- Moon, T.S., Y.J. Kim, J.R. Kim, J.H. Cha, D.H. Kim and C.M. Kim, 2009. Identification of process operating state with operational map in municipal wastewater treatment plant. *J. Environ. Manage.*, 90: 772-778.
- Oja, E., 1982. A simplified neuron model as a principal component analyzer. *J. Math. Biol.*, 15: 267-273.
- Oja, E., 1995. The nonlinear PCA learning rule and signal separation-mathematical analysis. Technical Report A 26, Helsinki University of Technology.
- Oja, E. and J. Karhunen, 1985. On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix. *J. Math. Anal. Appl.*, 106: 69-84.
- Runger, G.C. and F.B. Alt, 1996. Choosing principal components for multivariate statistical process control. *Commun. Stat-Theor. M.*, 25(5): 909-922.
- Tomita, R.K., S.W. Park and O.A.Z. Sotomayor, 2002. Analysis of activated sludge process using multivariate statistical tools-a PCA approach. *Chem. Eng. J.*, 90: 283-290.
- Wang, D. and J.A. Romagnoli, 2005. Robust multi-scale principal components analysis with applications to process monitoring. *J. Process Contr.*, 15(8): 869-882.
- Wold, S., K. Esbensen and P. Geladi, 1987. Principal component analysis. *Chemometr. Intell. Lab.*, 2: 37-52.
- Xu, L. and L. Yuille, 1995. Robust principal component analysis by self-organizing rules based on statistical physics approach. *IEEE T. Neural Networ.*, 6(1): 131-143.
- Yang, T.N. and S.D. Wang, 1999. Robust algorithms for principal component analysis. *Pattern Recogn. Lett.*, 20: 927-933.
- Yoo, C., P.A. Vanrolleghem and I.B. Lee, 2003. Nonlinear modeling and adaptive monitoring with fuzzy and multivariate statistical methods in biological wastewater treatment plants. *J. Biotechnol.*, 105: 135-163.
- Zhao, L.J., T.Y. Chai and Q.M. Cong, 2004. Multivariate statistical modeling and monitoring of SBR wastewater treatment using double moving window PCA. *Proceeding of the International Conference on Machine Learning and Cybernetics*, pp: 1371-1376.