ARTÍCULO

Combination of statistical process control (SPC) methods and classification strategies for situation assessment of batch process

Magda Ruiz, Joan Colomer, Joaquim Meléndez

Dept.de Electrònica, Informàtica i Automàtica
Universitat de Girona
Campus Montilivi, Edifici PIV
17071 Girona - Spain
{mlruizo,colomer,quimmel}@eia.udg.es

Abstract

The paper focuses on the development of a classification strategy to identify critic situation in batch process control. Data acquired from a batch execution is reduced by means of multiway principal component analysis in order to be assessed according to the statistical model of the process. Multiple situations have been categorized by a classification algorithm applied to the principal components in order to identify misbehaviour causes.

Palabras clave: Multiway Principal Component Analysis (MPCA), situation assessment, Batch Processes.

1. Introduction

Many strategies for fault detection and diagnosis are referenced in the bibliography. According to [18], fault diagnosis methods can been classify in three general categories: quantitative model based methods, qualitative model based methods and process history based methods, illustrated by figure 1.

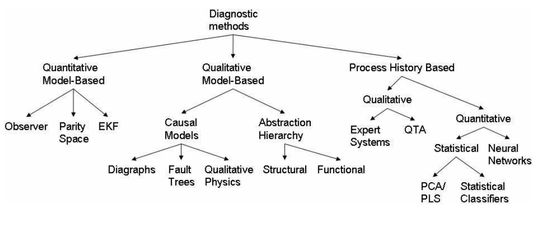

Figure 1. Classification of diagnostic algorithms, according to [18]

The solution proposed in this work falls in the third category; and particularly in the subgroup of statistical methods. A biological batch process for the treatment of wastewater has been used to develop and test the supervision method.

Multivariate Statistic Process Control (MSPC) methods have shown to be effective in detecting and diagnosing events that cause a significant change in the dynamic correlation structure among the process variables [3] some examples are: polymerization reactor process [12], pharmaceutical process [7], the elaboration at industrial scale of the polymer polypropylene oxide [20], WasteWater Treatment Plant [6] among others. These variables utilize the information directly and systematically and scientifically recognize the

ISSN: 1137-3601. ©AEPIA (http://www.aepia.org/revista)

normal operation behavior of the process. Different applications have been proposed in the literature according to this principle. In [2] a strategy to isolate sensors that are affected by nonconforming operation is described. It allows to distinguish between failed sensors and process upsets. In [4] MSPC is combined with wavelet properties, in this way was created adaptive multiscale MPCA in order to detect abnormal behaviors and to identify the major sources of process disturbances.

In this work a combination between MSPC and a classification tool is proposed. The combination of both methods improves the rsults obtained using only MSPC. The paper describes the operation of the SBR process in section 2. Then, section 3 is focused on those MSPC extensions for process monitoring. Section 4, the classification method is presented. And finally in section 5 and subsequent a example is presented and evaluated using data acquired from the real plant.

2. Biological batch process

A WasteWater Treatment Pilot Plant has been used in this work. The plant operates as a Batch Reactor (SBR) as Figure 2 depicts. In a SBR wastewater treatment plant nitrogen removal and elimination of organic matter is done with sludge. Sludge is responsible for the organic matter degradation and nitrogen removal. SBR Pilot plant is composed of a metal square reactor with a capacity of 200 liters of water to process. Wastewater is taken directly from a real station sited in Girona (Spain). Next, the wastewater is pumped to the reactor where the treatment is performed.

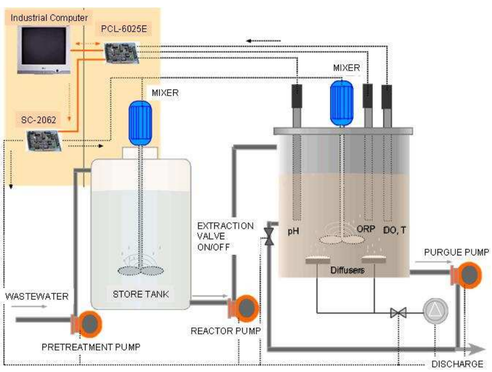

Figure 2. Real SBR pilot plant

In SBR pilot plant the nitrogen and organic matter are removed after a 8 hours cycle in which anoxic and aerobic stages are alternated. In an aerobic stage the ammonia is converted to nitrate and under anoxic condition nitrate is converted to nitrogen gas. Four process variables are monitored: pH, Oxidation Potential Reduction (ORP), Oxygen Dissolved (OD) and Temperature. The process is highly nonlinear, time-varying and subject to significant disturbances such as atmospheric changes, variation in the composition of influent. The process has been characterized statistically by its covariance matrix in order to study the correlation structure between variables and streams of them.

3. MSPC for batch processes

MSPC is a reduction technique based on classical statistical process control (SPC) theory extend to operate with multiple variables. Nowadays, it has also been adapted to characterise batch processes by considering as an additional dimension the number of batches (execution of a process according to a recipe) assuming the same length (same number of samples). The bases of MSPC for batch processes are the extensions of Principal Component Analysis (PCA) and Partial Least Squares (PLS) [11][6][10][5]. Extensions of principal component analysis are described in the next section.

3.1. Multiway principal component analysis (MPCA)

Consider a typical batch run in which j=1,2,...,J variables are measured at k=1,2,...,K time instants throughout the batch. Similar data will exist on a number of such batch runs i=1,2,...,I. All the data has been summarized in the $\underline{\mathbf{X}}$ (I x J x K) array illustrated in figure 3, where different batches are organized along the vertical side, the measurement variables along the horizontal side, and their time evolution occupant the third dimension. Each horizontal slice through this array is a (J x K) data matrix representing the time histories or trajectories for all variables of a single batch (i). Each vertical slice is an (I x J) matrix representing the values of all the variables for all batches at a common time interval (k) [10] [19].

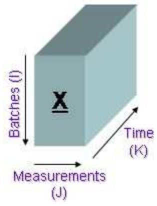

Figure 3. Arrangement of a three-way array $\underline{\mathbf{X}}$

MPCA is equivalent to performing ordinary PCA on a large two-dimensional $(2-D)$ matrix constructed by unfolding the three-way. Six ways of unfolding the three-way data matrix $\underline{\mathbf{X}}$ are possible [20]. In this work the unfolding $(IK$ x $J)$ in variable direction and $(I$ x $KJ)$ in batch direction are used. Undey and Cinar inspired in Wold [16] uses type $(IK$ x $J)$ (figure 4), motives within on-line monitoring of the batch process. The unfolding corresponds to type $(I$ x $KJ)$ is used by Nomikos and MacGregor [10] (figure 5), this unfolding is particularly meaningful because, by subtracting the mean of each column of this matrix $\mathbf{X}$, these procedures are subtracting the mean trajectory of each variable, thereby removing the main nonlinear and dynamic components in the data [9].

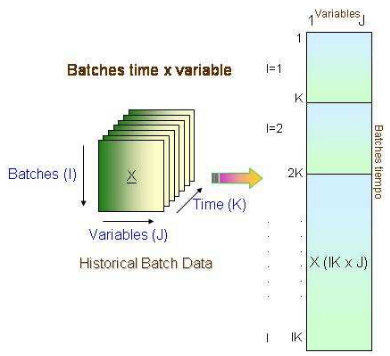

Figure 4. Decomposition of $\underline{\mathbf{X}}$ to 2-D $(IK$ x $J)$

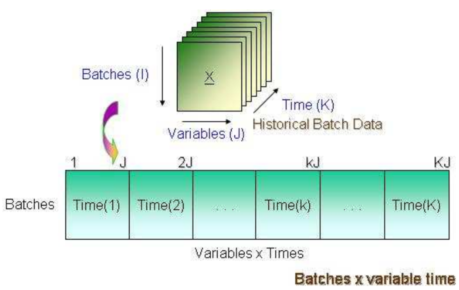

Figure 5. Decomposition of $\underline{\mathbf{X}}$ to 2-D $(I$ x $KJ)$

The objective of MPCA is to decompose the three-way $\underline{\mathbf{X}}$, into a large two-dimensional matrix $\mathbf{X}$. This decomposition is accomplished in accordance with the principal of PCA and separates the data in an optimal way into two parts: The noise or residual part $(\underline{E})$, which is as small as possible in a least squares sense, and the systematic part $(\sum_{r=1}^{R} t_r \bigotimes P_r)$, which expresses it as one fraction (t) related only to batches and a second fraction (P) related to variables and their time variation [10]. The MPCA algorithm derives directly from the NIPALS algorithm , resulting the matrix X. It is the product of score vector t_r and loading matrices P_r, plus a residual matrix E, that is minimized in a least-squares sense:

$$\underline{\mathbf{X}} = \sum_{r=1}^{R} t_r \bigotimes P_r \tag{1}$$

$$X = \sum_{r=1}^{R} t_r P_r^T + E = \hat{X} + E \tag{2}$$

MPCA decomposes the three-way $\underline{\mathbf{X}}$ array where $\bigotimes$ denotes the Kronecker product ($\underline{\mathbf{X}} = t \bigotimes P$ is $\underline{\mathbf{X}}(i,j,k) = t(i)P(j,k)$) and R denotes the number of principal components retained. The equation (1) is the 3-D decomposition while the equation (2) displays the more common 2-D decomposition [16].

3.2. Multiblock multiway principal component analysis (MMPCA)

In this case the data matrix $X(IxKJ)$ is divided into K blocks $(X_1, X_2, ..., X_K)$ in such a way that the variables from each time instant can be blocked in the same block (see figure 6) [6][16]. This approach has significant benefits because the latent variable structure is allowed to change at each phase in the batch processes. In the lower layer of the model, each data block is considered as a separate source of information and the details of the blocks are modelled by corresponding block model. In the upper layer, information from all blocks on the lower data level is combined and the relative importance of the different blocks, $X_{b.}$, for each dimension is obtained. In the upper layer information from the previous block, block scores $t_{b(k-1)}$, is combined with the block score vector from the lower layer [17][15].

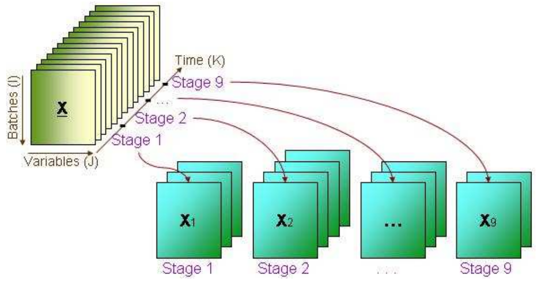

Figure 6. Dividing batch data into different phases

3.3. Control charts

Abnormal behavior of batch can be identified by projecting the batch onto the model. Control charts that are used in monitoring batch processes are generally based on the the Q-statistic and D-statistic, in which control limits are used to determine whether the process is in control or not.

The Q-statistic is a measure of the lack of fit with the established model. For batch number i, Q_i is calculated as:

$$Q_i = \sum_{j=1}^{J}\sum_{k=1}^{K}(e_{jk})^2 \sim gx^2_{(h)} \qquad (3)$$

where e_{jk} are the elements of E. Q_i indicates the distance between the actual values of the batch and the projected values onto the reduced space.

The D-statistic or Hotelling T^2 statistic, measures the degree to which data fit the calibration model:

$$D_i = t_i^T S^{-1} t_i \sim \frac{I(I-R)}{R(I^2-1)} F_{R,I-R} \qquad (4)$$

where S is the estimated covariance matrix of the scores. The D-statistic gives a measure of the Mahalanobis distance in the reduced space between of batch and the origin that designates the point with average batch process behavior.

4. Classification method

For classification, the Learning Algorithm for Multivariate Data Analysis (LAMDA) has been used [1]. This method takes advantage of hybrid logical connectives to perform a soft bounded classification.

LAMDA is proposed as a classification technique to apply to principal components selected for monitoring. The goal is to assess the actual situation according to profiles previously learned [1][8].

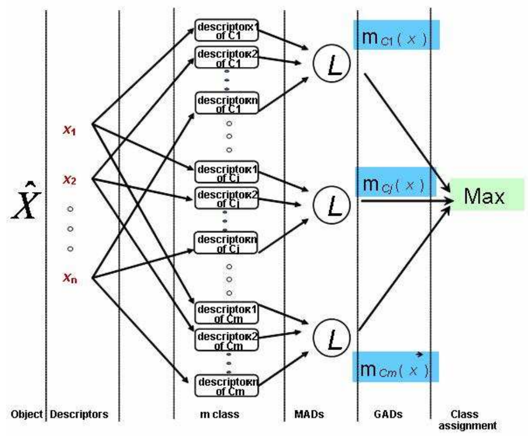

Figure 7. Basic LAMDA recognition methodology

Input data is presented to LAMDA as a set of observations or individuals characterized by its descriptors or attributes and recorded as rows. Principal components obtained in the MPCA step are used as input variables to be classified. Once, the descriptors are loaded, every individual is processed individually according to the desired goal [1]:

1. To classify the individuals according to a known and fixed set of classes.
2. To learn and adapt from a previous given set classes which can be modified according to the new individuals.
3. To discover and learn representative partitions in the training set.

The basic assignment of an individual to a class follows the procedure represented by figure 7. In this, MAD and GAD stand for Marginal (it takes into account only one attribute) and Global Adequacy Degree (obtained from the hybrid logical combination of the previously obtained MADs) respectively, of an individual to a given class. Equations (5) and (6) are used to calculate them. This classifying structure resembles that of a single neuron ANN [1].

$$MAD(d_ix_j/\rho_{i/k}) = \rho_{i/k}^{d_ix_j}(1-\rho_{i/k})^{1-d_ix_j} \quad (5)$$

where

d_ix_j = Descriptor i of the object
j $\rho_{i/k}$ =ρ of descriptor i and class k

$$GAD = \beta T(MAD) + (1-\beta)S(MAD) \quad (6)$$

Formalizing the description of LAMDA, it is possible to define an individual as a series of descriptors values d_1, ...,d_n such that each d_j takes values from the either finite or infinite set D_j. We will call universe or context to the Cartesian product $U = D_1$ x D_2 ... x D_j. Thus, any object or individual is represented as a vector $x = (x_1,...,$ $x_n)$ from U, such that each component x_j expresses the value for the descriptor d_j in the object x. The subset of U gathering all these vectors will be called data base or population. To assign individuals to classes MAD step will be calculated for each individual, every class and each descriptor, and these partial results will be aggregated in order to get the GAD of an individual to a class. The simplest way to build this system would be by using probability distributions functions, and aggregating them by the simple product, but that would force us to impose a series of hypothesis on the data distribution and independence which are too arbitrary. Finally, MAD and GAD have been used according to definitions of equation 5 and equation 6 respectively [1]. The hybrid connective used for GAD is a combination between a t-norm and a t-conorm by means of the β parameter. $\beta =$ 0 represents the intersection and $\beta = 1$ means the union. This parameter will -inversely- determine the exigency level of the classification, so it can be identified as a tolerance or exigency parameter.

5. Results

5.1. Types of batch process

The data obtained from the SBR process was analyzed under to points of view. The first one, based on analytical methods proposed in [13] where the sludge reaction is explained. The second one, was a preliminary MSPC analysis where some batches are detected to be outside the control limit. This study create five types of SBR batch process: Electrical fault, variation composition, atmospheric changes (corresponding to rain), equipment defects and normal behavior. According to the classification it is possible to quantify the number of batches for each group, in the Table 1 all batches of the SBR process are summarised. There are 60 (equivalent to 33,5 %) batches with abnormal behavior. The normal behavior was the most common type (66,5 %) with a higher nitrogen efficiency than legally required effluent standards, which are classified according to the final quality of the wastewater.

Type of batch process	Quantity	%
Atmospheric changes	17	9.50
Equipment defects	8	4.47
Variations in the composition	33	18.44
Electrical Fault	2	1.12
Normal behavior	119	66.48

Table 1. Batch classification by group

5.1.1. MPCA: batch direction

Each batch lasts 8 hours (5760 samples for each variable sampled every 5 seconds). Only 392 samples from each one of the four acquired variables have been used in order to reduce computational cost resulting a $IxKxJ$ = 179x4x392 array, ($\underline{\mathbf{X}}$) for the collection of 179 available batches. MPCA algorithm was applied to the three-way data array, $\underline{\mathbf{X}}$ unfolded in the batch direction (I x KJ) resulting 8 principal component. So, the new dimensionality becomes (179 x 8). The statistical model was created with eight components, which explain 92,79 % of the total variability. To examine the process data in the reduced projection spaces (defined by a small number of latent variables), the variables contribution analysis are made; as is shown in Figure 8 the temperature variable is positively correlated with loadings 1 where can be appreciate that in sample 1113 had a increase of the temperature. From the Figure 9 represents loadings 2 where Load2 represents at ORP variable.

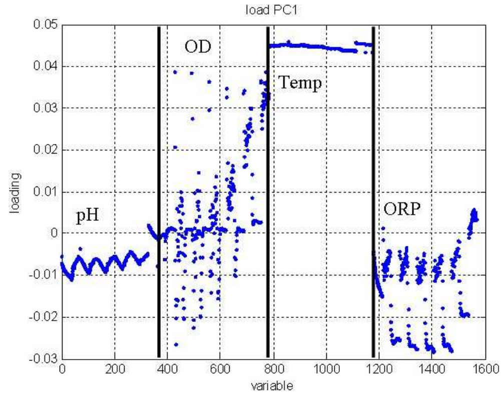

Figure 8. Variable loadings for the principal component

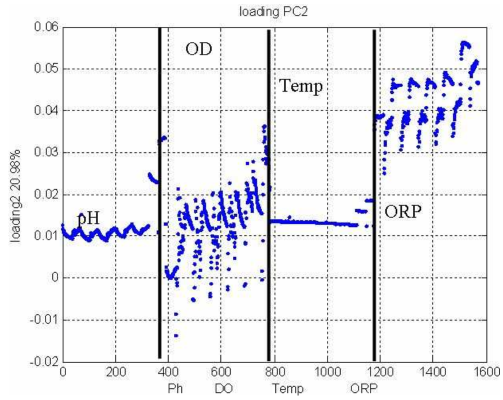

Figure 9. Variable loadings for the second component

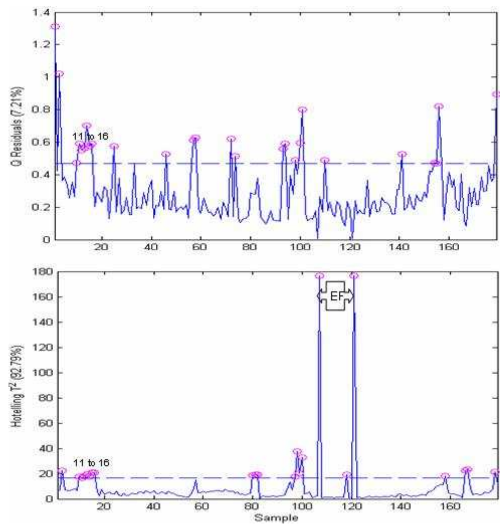

Figure 10. Multiway PCA. Q and T^2 charts with 92,79 % confidence limits

Figure 10 shows the Q and T^2 charts for all process batches. In the Q chart, it can be seen that some batches exceed its limits. These batches have several behaviors. In T^2, two batches are outside. These batches had electrical fault (EF).

In Table 2 the batches outside the model are summarised. In the Q chart, only a third of the total the abnormal behavior is detected, furthermore there are 8 false alarms. The T^2 chart has 20 batches with abnormal behavior (without false alarm). 39 about 60 of the abnormal behavior can be detected, 9 batches are in both charts.

Type of batch process	Q		T	
	Quantity	%	Quantity	%
Atmospheric changes	9	5.03	4	2.23
Equipment defects	0	0.00	6	3.35
Variations in the composition	11	6.15	8	4.47
Electrical Fault	0	0.00	2	1.12
Normal behavior	8	4.47	0	0.00

Table 2. MPCA classification

5.1.2. MPCA: variable direction

Three-data matrix $\underline{\mathbf{X}}$ has been unfolded in variable direction too (IK x J). The model was developed with dimensions (70168 x 4), where MPCA squeezes in 3 principal components explaining the 95,18 % of the total variability. In figure 11 a projection on the first and second component plane of the statistical model.

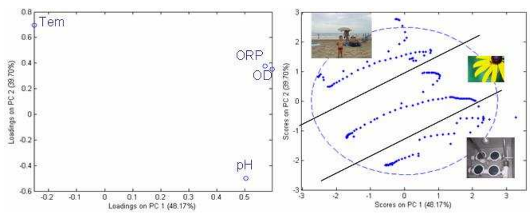

Figure 11. Variables weighs and model in variable direction

The batches are sequentially ordered and there are 3 sections into the model. Each section corresponds to batch gathered during specific seasons. Test of SBR process match the first month with monitoring; Spring is the batches developed in spring season finally summer correspond to cycles in summer season. Temperature contribution was demonstrated to be less important than others variables (−0,25 of first component) in consequence, it was omitted and a new model was built using only 3 variables. Figure 12 shows how the the new model is equally representative.

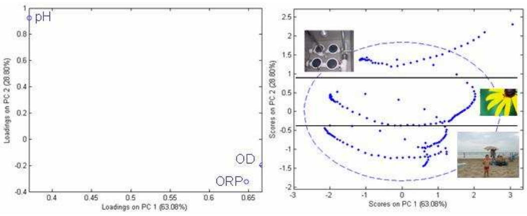

Figure 12. Variables weighs and model in variable direction without Temperature variable

5.1.3. Multiblock MPCA

The SBR pilot plant consists of 6 stages in which the latent variable structure can change due to different environments. Applying Multiway MPCA, the data matrix $\underline{\mathbf{X}}$ can be break. In this way, it is possible to work with the total three-way data array, $\underline{\mathbf{X}}$, with dimensions 179 x 4 x 5760. Data array for each stage are: cycle 1 (179 x 4 x 780);cycle 2 (179 x 4 x 780);cycle 3 (179 x 4 x 780);4 (179 x 4 x 780);cycle 5 (179 x 4 x 780);cycle 6 (179 x 4 x 804);purge cycle (179 x 4 x 36);settling cycle (179 x 4 x 720);draw cycle (179 x 4 x 300). Using the control charts by each stage, it is possible to observe the following: Batches 11 to 17 have variation in the composition and these batches are identified by the Q and T^2 control charts. The alarms by each stage are summarized in Table 3 (common batches are discounted). Purge, settling and draw are stages without nitrogen removal, they have more false alarms than other stages.

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Purge	Settling	Draw
outside	26	31	34	26	28	21	16	20	35
false alarm	8	3	5	3	4	1	5	8	23

Table 3. Alarms by each stage

These Multiblock charts supply knowledge by stages which potentially help to fault location and diagnosis. Furthermore the data interpretation is easier. Some types of batch process with large duration fault have been found to be present in the 6 stages models, for example the batches 10 to 16.

5.1.4. Conclusions of MSPC

Initially with the combination of MPCA and analytical methods it was possible to classify all batches. Individually, the model developed with MPCA in batch direction has produced satisfactory results because knowledge of the process was obtained while the model developed with MPCA in variable direction allows to detect the relationsship between process behavior and environment (rain period, summer, among others) and Multiblock MPCA gives detail of the process. In general, MSPC has been used to detect abnormal behavior in SBR process, by projecting the data into a lower dimensional space that accurately characterizes the state of the process. The use of a classification tool to the new variables allows a simple identification and grouping of similar situations according to a matching criteria.

5.2. Classification for situation assessment

Initially, MPLS was used. This technique is a dimensionality reduction that maximises the relation between the matrix X (I x JK) and the predicted matrix Y [14] (179 x 5) where 179 is the number of historical data batches and 5 are the types of batch process. The model make did not describe the process because matrix Y was created with the results obtained of the preliminary MSPC analysis. Matrix Y should be constructed with quality variables which are obtained each three days, finding now the missing problem. Thus, a classification tool for situation assessment is used: MPCA + classification tool.

5.3. MPCA classification

$\hat{X}$ is the principal components by each batch with dimensions 8 x 179. These were sued as descriptors to feed into LAMDA algorithm to discover relevant classes under an unsupervised schema. The tool automatically classified the data in eleven classes (11). Table 4 compares the classes and the types of batch process. According to these results, it is possible to identify classes that only contain batches with equipment defects, electrical faults, atmospheric changes and variation in the composition. The classes 1,9 and 10 correspond to normal behavior. The group 6 is associated to atmospheric changes. Classes 3 and 11 represent variations in the composition while classes 7 and 8 include electrical fault. Finally, the classes 2, 4 and 5 groups different types of batches. The predominant class (class 1) has 48,04 % of the total historical data, this class represents the normal behavior. The class 5 is abnormal behavior due to atmospheric changes and equipment defects.

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	Class 10	Class 11
normal	79	17	0	8	0	0	0	0	17	4	0
atmospheric changes	2	1	0	10	3	1	0	0	0	0	0
equipment defects	1	3	0	0	2	0	0	0	0	0	2
variation composition	4	11	7	1	0	0	0	0	3	1	0
Electrical Fault	0	0	0	0	0	0	1	1	0	0	0
total	86	32	7	19	5	1	1	1	20	5	2

Table 4. Composition by class

The relationship between the class and principal components is another observation. The 8^{th} component is less predominant because it does not change. It indicates that $\hat{X}$ can be computed using only seven descriptors. Then, the total variability will be 90,54 %. Consequently, only 7 principal components are used in the analysis Multiblock MPCA.

5.4. MMPCA classification

According to previous analysis, there are seven principal components (seven descriptors for each stage). Classification tool is used individually at every stage taking the whole set of batches. It resulted that at different stages the numbers of classes was very also different. Likewise to MPCA classification, the classes were marked (Table 5). Table 6 summarises the error for this classification. Other observation: electrical Fault is present in cycles 2 and 6 because one batch experimented a sags in two cycles (Remember Table 1). Types of normal behavior are the classes more populated.

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9
Electrical Fault		11				7,8			
Variation in the compositio	3,6,7	3,5,19	2,5,6,17, 18	2,6,16,17	2,3,7,17, 18	2,4	2,4,5,6,9, 15,16	2,5,14	2
Equipment defects	9,18	8	8	9	12	12			
Atmospheric changes	15	9,10,14	11,14	11,12,14, 19	10,11,19, 21	6	11	7,8,11,18, 19	8
Normal behaviour	1,2,4,5,8, 10,11,13, 14,17,21, 22, 23,24,25	2,4,7,12, 15, 17,18	1,3,4,7,12 13,15,16, 19	1,3,4,5,7,8 , 10,13,18	1,4,5,6,8,9 , 13,15,16	1,3,9,11	1,7,8,10, 13, 14	1,3,4,6,9, 10,12,13, 15,17	1,3,5,6,7,9 , 10,11,12, 13,15,16, 19
Not classified	2,16,19,20	1,6,13,16	9, 10	15	14, 20	5, 10	3,12	16	4,14,17,18

Table 5. Classes by each cycle

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6	Stage 7	Stage 8	Stage 9
Quantity not classified	20	29	9	6	9	8	9	6	47
% not classified	11,17%	16,20%	5,03%	3,35%	5,03%	4,47%	5,03%	3,35%	26,26%
Batch wrongly classified	4	8	7	9	8	13	12	10	14
% mistake	2,23%	4,47%	3,91%	5,03%	4,47%	7,26%	6,70%	5,59%	7,82%

Table 6. Classes by each cycle

6. Conclusions

Multivariate Statistical Process Control has been used to detect abnormal behavior in SBR process by projecting the data into a lower dimensional space that accurately characterizes the state of the process. Therefore, the new variable matrix is smaller. The use of a classification tools has been teste with previously known data to verify the utility of it to discover clusters of data in the historical registers useful for further situation assessment. MSPC and classification tool. Splitting data into meaningful groups allows a faster localization and identification of faults reporting similar experiences.

In order to improve the results and to process the data faster, it is necessary to developed a technique that combine the dimensionality reduction and nonlinear classification instead of the classical strategy. The use of a classification tool to the new variables allows a simple identification and grouping of similar situations according to a matching criteria.

Acknowledgement

This work is part of the research project Development of a intelligent control system apply to a Sequencing Batch Reactor by loads (SBR) for the elimination of organic matter, nitrogen and phosphorus DPI2005-08922-C02-02 supported by the Spanish Government and the FEDER Founds.

Referencias

[1] J. Aguliar-Martin and R. Lopez. The process of classification and learning the meaning of linguistic descriptors of concepts. *Approximate Reasoning in Decision Analysis*, pages 165–175, 1982.

[2] Fuat Doymaz, Jose Romagnoli, and Ahmet Palazoglu. A strategy for detection and isolation of sensor failures and process upsets. *Chemometrics and Intelligent Laboratory Systems*, 55:109–123, 2000.

[3] Alberto Ferrer, editor. *Control Estadistico MegaVariante para los Procesos del Siglo XXI*. 27 Congreso Nacional de Estadistica e Investigacion Operativa (Spain), 2003.

[4] Dae Sung Lee, Jong Moon Park, and Peter Vanrolleghem. Adaptive multiscale principal component analysis for on-line monitoring of a sequencing batch reactor. *Journal of Biotechnology*, 116:195–210, 2005.

[5] Dae Sung Lee and Peter A. Vanrolleghem. Adaptive consensus principal component analysis for on-line batch process monitoring. Technical report, Fund for Scientific Reseach - Flander (F.W.O.) and the Ghent University Resarch Fund, Coupure Links 653, B-9000 Gent, Belgium, 2003.

[6] Dae Sung Lee and Peter A. Vanrolleghem. Monitoring of a sequencing batch reactor using adaptive multiblock principal component analysis. *Biotechnology and Bioengineering*, 82(4):489–497, mai 2003.

[7] J.A. Lopes, J.C. Menezes, J.A. Westerhuis, and A.K. Smilde. Multiblock pls analysis of an industrial pharmaceutical process. *Biotechnol Bioeng*, (80):419–427, 2002.

[8] K. Moore. Using neural nets to analyse qualitative data. *A Marketing Research*, 7(1):35–39, 1995.

[9] Paul Nomikos and John MacGregor. Multivariate spc charts for monitoring batch process. *Technometrics*, 37(1):41–59, feb 1995.

[10] Paul Nomikos and John F. MacGregor. Monitoring batch processes using multiway principal component analysis. *AIChE*, 40(8):1361–1375, aug 1994.

[11] Paul Nomikos and John F. MacGregor. Multi-way partial least squares in monitoring batch processes. *First International Chemometrics InterNet Conference*, 1994.

[12] Aras Norvilas, Eric Tatara, Antoine Negiz, Jeffrey DeCicco, and Ali Cinar, editors. *Monitoring and fault diagnosis of a polymerization reactor by interfacing knowledge based and multivariate SPM tools*, number 0-7803-453. American Control Conference, 1998.

[13] S. Puig, M.T. Vives, Ll. Corominas, M.D. Balaguer, and J. Colprim, editors. *Wastewater nitrogen removal in SBRs, applying a step-feed strategy: From Lab-Scale to pilot plant operation*. 3ª IWA Specialised Conference on Sequencing Batch Reactor, Australia, feb 2004.

[14] Evan L. Russell, Leo H. Chiang, and Richard D. Braatz. *Data-driven techniques for fault detection and diagnosis in chemical processes "Advances in Industrial Control"*. ISBN 1-85233-258-1, London, 2000.

[15] Age K. Smilde, Johan A. Westerhuis, and Ricard Boqué. Multiway multiblock component and covariates regression models. *Journal of Chemometrics*, 14:301–331, 2000.

[16] Cenk Undey and Ali Cinar. Statistical monitoring of multistage, multiphase batch processes. *IEEE Control Systems Magazine*, 22(5):40–52, oct 2002.

[17] Cenk Undey, Bruce A. Williams, and Ali Cinar, editors. *Monitoring of Batch Pharmaceutical Fermentations: Data Synchronization, Landmark Alignment, and Real-Time Monitoring*. 15th Triennial World Congress, Barcelona Spain 2002 IFAC, 2002.

[18] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. Kavuri. A review of process fault detection and diagnosis part i quantitative model-based methods. *Computer and Chemical Engineering*, 27:293–311, 2003.

[19] Johan A. Westerhuis, Theodora Kourti, and John F. MacGregor. Analysis of multiblock and hierarchical pca and pls models. *Journal of Chemometrics*, 12:301–321, 1998.

[20] Manuel Zarzo and Alberto Ferrer. Batch process diagnosis: Pls with variable selection versus block-wise pcr. *Chemometrics and intelligent laboratory systems*, 73:15–27, jun 2004.