

Adaptive threshold PCA for fault detection and isolation

Marco A. Márquez-Vera^{1*}, Omar López-Ortega², Luis. E. Ramos-Velasco³, Andrea Rodríguez-Romero⁴, Julio C. Ramos-Fernández⁵, Jorge A. Hernández-Salazar⁶

^{1,4,5}Mechatronics and Biotechnology Engineering, Polytechnic University of Pachuca, Zempoala Hgo; Mexico

²Electronics and Computation Academy, Autonomous University of Hidalgo State, Hidalgo, Mexico

³Aeronautics Engineering, Metropolitan Polytechnic University of Hidalgo, Tolcayuca, Hgo. Mexico

⁶Computer Systems, National Technological Institute of Mexico, Huejutla, Hgo. Mexico

Emails: ¹marquez@upp.edu.mx, ²lopezo@uaeh.edu.mx, ³ramos@upmh.edu.mx,

⁴andrearodriguezayo@gmail.com, ⁵jramos@upp.edu.mx, ⁶acad_huejutla@tecnm.mx

*Corresponding Author

Abstract—Fault diagnosis is an important issue in industrial processes to avoid economic losses, process damage, and to guarantee safe working conditions for the operators. For high scale industrial processes the data-driven based methods are the best solution for process monitoring and fault diagnosis. Thus, in this paper, the principal component analysis is shown to detect and isolate faults. Also, a dynamic threshold is implemented to avoid false alarms because incipient faults are difficult to be detected. As a case of study, the Tennessee Eastman (TE) process is used to apply this strategy because the interaction among five units with internal control loops makes difficult to have an approached model. As results are shown the detection times, for cases where were analyzed incipient faults, the time required for fault detection must be improved, in this work, an adaptive threshold was used to reduce the false alarms but it also increases the detection times. It was concluded that the Q chart gave a better result for fault detection; the isolation times were similar to the detection ones. Two incipient faults could not be detected, the fault detection rate was similar to the shown in literature, but the detection times were better in 35% of the cases, unfortunately for four faults the detection times were bigger than the reported in other papers. It is proposed to help this method with independent component analysis due it is not guaranteed to have a Gaussian distribution in the samples.

Keywords—Fault detection, PCA, incipient faults

I. INTRODUCTION

In industrial processes the variables monitoring is a common way to determine the process healthily, nowadays it is common to have a high level of automation and data acquisition systems like the SCADA; in this way, the historical data are available. Thus, a data-based method for fault detection and isolation (FDI) is a good option; the common strategies use statistical analysis [1-3] to evaluate the signals to look for anomalous behavior.

An approached model can be difficult to propose for large scale industrial processes as in the conclusions of [2] was indicated. In this way, the Hotelling statistic T^2 and the square prediction error (SPE) are used here for FDI [3].

The fault detection consists of awareness regarding the occurrence of a problem in the process [1]. And the recognition of the event that occurred is the fault isolation [4].

When both activities are developed is common to find the concept of fault diagnosis [5].

A statistic method for FDI commonly used is the principal component analysis (PCA), for example, in [6] is employed the PCA to detect anomalies in synchronous generators by using a threshold obtained with the T^2 chart. Hierarchical PCA was used for fault detection due to sensors degradation in [7]. Also, a combination between ANFIS and PCA can be found in [8] for fault diagnosis in the Tennessee Eastman process. Multiscale PCA-based fault detection for the same process was shown in [9].

The Tennessee Eastman (TE) process is an industrial process originally developed in [10] to evaluate fault-tolerant control strategies and FDI methodologies. This benchmark is widely used due to the number of variables and faults that can be simulated, for this process is difficult to propose an approach model to have a model-based FDI. This chemical process has five principal units: a chemical reactor, a product condenser, a recycle compressor, a liquid-vapor separator, and a product stripper [11]. The TE process has 41 variables that can be measured, 11 input variables, and it is possible to simulate 20 different faults. The sample time is three minutes for the first 22 outputs named $x(1)$, $x(2)$, ... $x(22)$, for the compound variables $x(23)$ - $x(36)$ the measures are obtained after six minutes, and for the compound variables $x(37)$ - $x(41)$ the sampling time is 15 minutes [12].

A general scheme of the TE process can be appreciated in [11], where it is also possible to see a description of the variables and faults that describes the process. The faults are denoted as IDV1, IDV2, ..., IDV20. Some of them are particularly difficult to identify because they are imperceptible by regarding the output signals, thus they are named incipient faults [12]. A revision of the TE process model was made in [13] because of inconsistencies in the ODE solver.

Due to the difficulty of proposing an approached model for the TE process, data-driven methods are used for this kind of system. Almost all applications using PCA in the TE process are focused on fault detection and dimensionality reduction, these works show the fault detection rate (FDT) as



the criterion to evaluate the algorithm performance [1, 2, 3, 8, 13, 14]. Here is proposed to evaluate the number of samples needed to detect a fault because it is important to recognize an anomalous behavior as soon as possible to avoid the fault evolution [15].

In this paper, an adaptive threshold sensitive to the variance [16] was implemented to reduce false alarms and the fault isolation was obtained using the PCA as well to get the FDI in the Tennessee Eastman process, evaluating the performance of this proposes using the fault detection rate [17] and the detection times required. It is important to note that detecting a fault is not sufficient because the operator must know the problem presented in the process. Thus, the PCA-based fault isolation for the TE process, here presented, is the main contribution.

II. METHODS

There are 41 variables to be processed from the TE process, and it can be assumed that at least one of them will highlight a fault present in the monitored process as a hypothesis. In order to improve the signal processing to look for a fault, it is possible to select the most representative variables. The PCA can be used to reduce the amount of data [16] and is useful to deal with data that contain a high degree of correlation between variables as in [9] was explained.

The dataset used to generate the PCA model was obtained by simulating the TE process in normal operation conditions, the dataset contains 4801 samples. For the faulty operations, a fault was introduced between the 1000th and 2000th samples. The faults are introduced into the model by assigning a 1 into a vector of length 20 with the other elements equal to zero [13]; this vector is connected to the disturbance input in the model. The fault detection by using PCA was developed in-line using the PCA model obtained in normal conditions [14].

The data measured from the TE process are noising, assuming that the noise has Gaussian distribution; the PCA can be used also to reduce the noise magnitude. For the PCA implementation is required the condition of normalizing the measured variables grouped in the matrix $X \in R^{n \times m}$ to have a zero mean value and unitary variance in the set \bar{X} [17], because different process variables have different units and mean values. Here, n is the number of samples and m is the number of variables.

A. PCA FAULT DETECTION

Once the data were normalized in \bar{X} , the covariance matrix $\Sigma \in R^{m \times m}$ is computed to know which variables move in the same direction [2], this matrix is computed using

$$\Sigma = \frac{1}{1+n} \bar{X}^T \bar{X}, \quad (1)$$

where n is the number of samples to normalize the covariance matrix. The PCA is used to get a reduced dimension space to reduce redundancy between variables, for this purpose is necessary compute eigenvector decomposition for Σ . One way to obtain the eigenvectors is to apply the singular value decomposition (SVD) [14] denoted by

$$\Sigma = V \Lambda V^T \quad (2)$$

where Λ is a diagonal matrix whose elements are the eigenvalues of Σ in descendent order and V is a matrix whose columns are eigenvectors of Σ [17].

Now, it is possible to select some eigenvalues to approach the data set reducing redundancy between variables, i.e.

$$\Sigma \approx \hat{\Sigma} = S P^T \quad (3)$$

where P is a matrix that has the eigenvectors selected and S is a matrix called the score matrix. The dimensions of these matrices depend on how many eigenvalues are considered to build $\hat{\Sigma}$, now some ways to select the eigenvalues (principal components) are described.

- One way to select the number of principal components is by taking the first eigenvalues that have almost all the information; thereby, the ones that have more correlation are ignored. To select the principal components most representative [3] is widely used the cumulative percentage variance (CPV) to have 85% of relevant information as in [16] is mentioned. The CPV is calculated using

$$\text{CPV}(a) = \frac{\sum_{i=1}^a \lambda_i}{\text{trace}(\Sigma)} \times 100, \quad (4)$$

where λ_i is the i -th element in the diagonal of Λ and a is the number of principal components to be selected.

It was obtained that 33 principal components are required to have the 85% of the information for fault detection, these 33 principal components are obtained using only the first 33 eigenvectors (columns) of matrix V from the SVD; thus, in this case, the loading matrix would be $P \in R^{41 \times 33}$. The principal components are obtained by a linear transformation using a new basis [2], it means that they are not a specific measured variable from X , the remaining 8 eigenvalues are discarded because of their correlation.

- Another possibility for dimension reduction is to use parallel analysis (PA), to get a, now in this way $a = 12$ as shown in Fig. 1, this result is also shown in the Master Thesis [18] in its Fig 4.3.

The selection of a was by using the measures without any fault, and 4801 samples were used for each available variable by using the loading matrix $P \in R^{41 \times 12}$ which has the first a columns of V , being the score matrix $S \in R^{4801 \times 12}$ defined as

$$S = \bar{X} P, \quad (5)$$

a new data set \hat{X} is obtained, where the correlations and linear dependencies among variables were removed. The residual subspace is obtained with the last $m-a$ principal components.

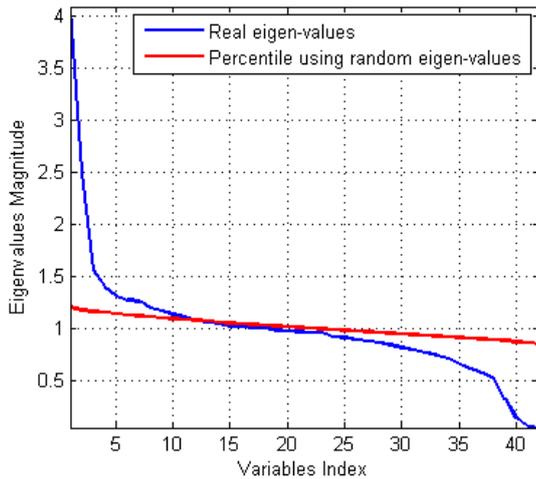
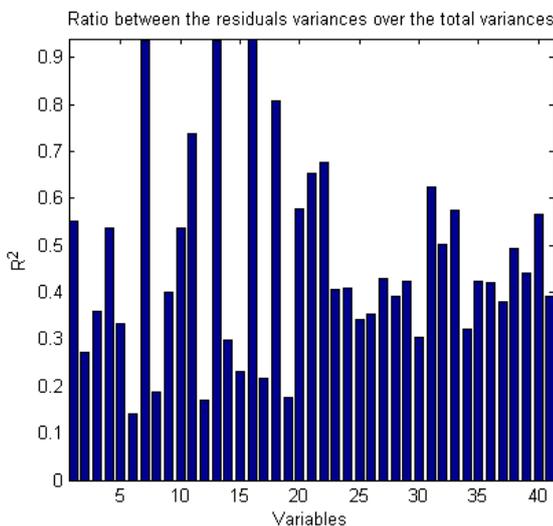


Fig. 1. Parallel analysis [18].

By using only a principal components, the PCA model reflects the main structure of the original data, capturing the maximum variance to prevent the loss of information [19]. The accuracy of the model can be estimated using the ratio between the residual variance over the total variance denoted by (6). In the case of using all the principal components $R_a^2 = 1$, but when only some of them are selected, this value decreases.

$$R_a^2 = 1 - \frac{\text{var}(\bar{X} - \hat{X})}{\text{var}(\bar{X})}, \quad (6)$$

here, a represents that only a principal component were used, and this value is computed for each variable (column in \bar{X}). In Fig. 2 are shown the ratios obtained.

Fig. 2. R_a^2 to evaluate the PCA model for each variable.

Once the PCA model was obtained, the data \bar{X} can be decomposed as (7)

$$\bar{X} = \hat{X} + \tilde{X}, \quad (7)$$

where $\hat{X} = \bar{X}PP^T$ and $\tilde{X} = \bar{X}\tilde{P}\tilde{P}^T$ are the modeled and unmodeled variations respectively, and the matrix $\tilde{P} \in$

$R^{41 \times 29}$ has the columns from $a+1$ to the last of matrix V that has the eigenvectors of the correlation matrix Σ .

Thereby, these two parts are necessary to evaluate the variables behavior for fault detection. The modeled variations monitoring can be implemented by using the Hotelling chart T^2 [2] computed by (8). The unmodeled behavior is evaluated with the Q statistic, also called the square prediction error (SPE) [16]. Now, the detection indexes for these statistics are calculated in the following way:

- The T^2 statistic is obtained using (8)

$$T^2 = \bar{X}P, \Lambda_a^{-1}P^T\bar{X}^T, \quad (8)$$

where $\Lambda_a \in R^{12 \times 12}$ has the first a columns of matrix Λ , $P \in R^{41 \times 12}$ is the loading matrix, and $\bar{X} \in R^{4801 \times 12}$ is the normalization of data. Now, a threshold T_a^2 is calculated using the F distribution [19] using (9),

$$T_a^2 = \frac{(n^2 - 1)\alpha}{n(n - \alpha)} F_a(\alpha, n - \alpha), \quad (9)$$

being $F_a(\alpha, n - \alpha) = 2.3$ for $\alpha = 0.05$.

- The Q chart the is computed with (10)

$$Q = \bar{X}\tilde{P}\tilde{P}^T\bar{X}^T \quad (10)$$

From chart Q , the threshold Q_α in used for fault detection, being obtained using (11)

$$Q_\alpha = \theta_1 \left(1 + \frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right)^{\frac{1}{h_0}} \quad (11)$$

where $\theta_i = \sum_{j=a+1}^m \lambda_j^i$ and $h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$, here, λ_i is the i -th element of the diagonal matrix Λ , and $c_\alpha = 1.96$ was obtained from the normal distribution with a level of significance $1 - \alpha = 0.95$ and $\gamma = 0.05$ [20].

It is important to remark that at least a certain amount of data should be acquired to compute a relevant T^2 or Q charts to propose a threshold to detect a fault. It could be noted that if the number of features to obtain a model grows, the amount of data required grows exponentially; the PCA can be used to reduce the dimensionality of the data set. The TE process has 41 variables, and an exploratory data analysis needs to identify the relationships between features from $41(41-1)/2=820$ cases. In areas like adaptive control and parameter identification, a good option is to have 10 times more measures than parameters to obtain, thus 410 samples can be used to develop a PCA model.

An interesting application of PCA to monitor the TE process was developed in [20], where was used a moving window to detect a fault, in his Fig. 4 is shown the robustness depending on the amount of data used, it is clear that by using more data, the robustness is improved, the best result was obtained with 260 samples. In this work, the simulation for the PCA application had 4801 samples to get a PCA model, and the faults were applied during 500 or 1000 samples.

A more important aspect to apply PCA is the threshold used for fault detection, in this way, in [21] is explained that

two conditions must be considered to have detectability; firstly, the projection of fault data, made with the T^2 or Q charts, does not vanish; and it also will be considered that the fault data projection must exceed the threshold T_{α}^2 or Q_{α} . In the TE model benchmark used, the faults are applied by writing one into an input vector, zero means that this particular fault will not be introduced [13]. The magnitude faults are not controlled, but some of them are difficult to detect because their effect is not visible with simple inspection, these kinds of faults are called incipient.

As a first result for fault detection, the data projection to the principal components subspace \hat{X} was used to compute the T^2 chart, this statistic is shown in the Fig. 3 in blue color. When the fault IDV2 is applied in the sample 1000, the T^2 , obtained from data projected using 12 principal components, is represented by the red line; thus, a fault can be inferred when the new statistic is bigger than the obtained with data in normal operation. In this case, the fault IDV2 was detected after 154 samples and it could be said that this fault is easy to detect. For other cases, it could not be detected the faults by observing their statistics; therefore, a threshold can be used for fault detection as in [18] is mentioned.

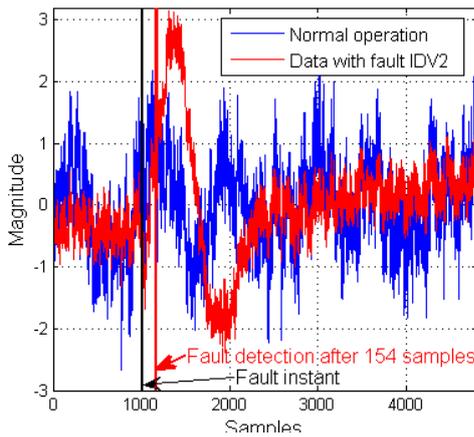


Fig. 3. PCA when IDV2 takes place.

In the Fig. 4 is possible to see that the measured variables are noisy, the PCA model has gaussian noise in the output signals with a standard deviation commonly found in sampled variables [10].

This method cannot always detect all the possible faults in the model, as in Table 11 of [2] was shown. The faults IDV3, IDV5, IDV14, IDV15, IDV16 and IDV19 almost never were detected using the PCA, in Fig. 5 are shown the T^2 and Q charts for the case of having the fault IDV19 at sample 1000, and it can be inferred that there are false positives and false negatives; therefore, this fault cannot be detected correctly.

In [2] a modified PCA was presented for fault detection in the TE process, where if $Q > Q_{\alpha}$ or $T^2 > T_{\alpha}^2$, then a fault has been detected. The condition for this technique application is that the fault must appear, at least, 40 samples after the simulation begins. The T^2 and Q charts obtained when fault IDV1 appears at sample 1000 are shown in Fig. 4 and it is possible to detect this particular fault; however, some

faulty samples can be omitted having in this way false positives but no false alarms.

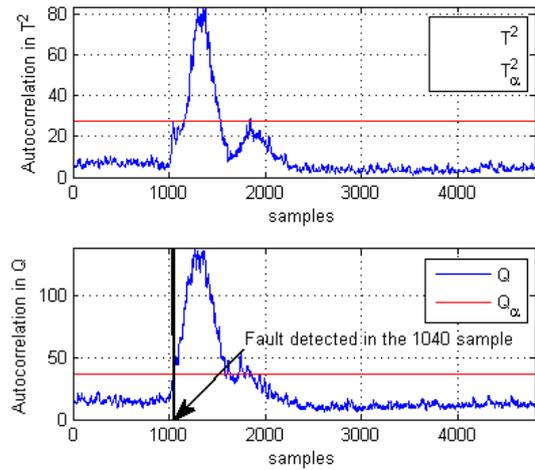


Fig. 4. T^2 and Q charts.

An adaptive threshold can avoid false alarms (FA), one way to reduce the FA rate is by using a low-pass filter [22], but the filter increases the number of samples needed to detect each fault. In Table 3 of [23] was shown the result of implementing the independent component analysis (ICA) plus a wavelet filter for denoising the signals in the TE process, where the more complicated faults to be detected were IDV3, IDV9, and IDV15 as also is indicated in [1].

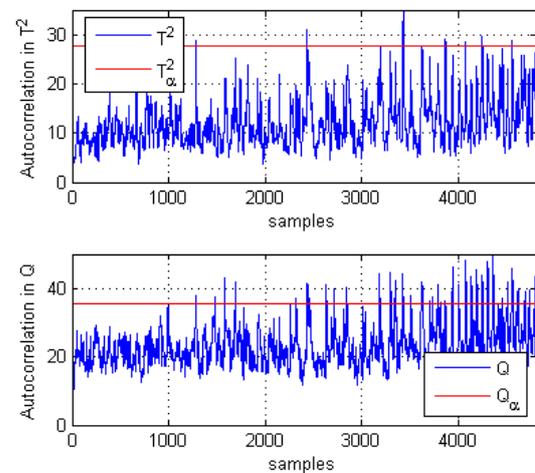


Fig. 5. T^2 and Q charts for fault IDV19.

In [16] a variable threshold was proposed for dynamic fault detection. This threshold is adapted according to the data variance by using the average (12) and the variance of T^2 (13)

$$\mu = \frac{1}{n} \sum_{i=1}^n T_i^2, \tag{12}$$

$$\sigma = \frac{1}{n+1} \sum_{i=1}^n (T_i^2 - \mu^2). \tag{13}$$

With a significance level $(1-\alpha)$, the adaptive threshold T_{adp} is computed by using (14)

$$T_{adp} = \mu(1 + z), \tag{14}$$

where typically $\alpha = 95\%$ and z is a coefficient related to a such that for the probability of having the mean μ with the significance level $(1-\alpha)$, is necessary that $\text{prob}(-z < \mu < z) = 1-\alpha$. For this condition z can be proposed using a cumulative distribution function, in [16] was defined the function (15)

$$Y(z) = \text{prob}(\mu < z) = 1 - \frac{\alpha}{2}, \quad (15)$$

from [19] was proposed to use $z = 1.5$ for having $\alpha = 0.95$.

The T_{vsa} threshold defined by (16) is sensitive to the data variance which is important because each variable has its own variance:

$$T_{vsa} = \begin{cases} T_{\alpha}^2, & \text{if } T_{\sigma}^2 \geq T_{adp}, \\ \mu + \sigma z, & \text{if } T_{\alpha}^2 < T_{adp} \text{ and } \mu \leq \sigma \\ \mu(1 + z), & \text{if } T_{\alpha}^2 < T_{adp} \text{ and } \mu > \sigma \end{cases} \quad (16)$$

In this way T_{vsa} uses the peaks of the signal to avoid false alarms, which is important in the transient-state due to changes in the system behavior. This sensitive threshold can be appreciated in Fig. 6 for the TE process simulation without any fault, a detail is shown to see the threshold adjustment.

Here the idea presented in [16] was used; the number of samples needed to detect each fault is shown in Table 1. Also a robust PCA can be implemented in order to reduce false alarms [24].

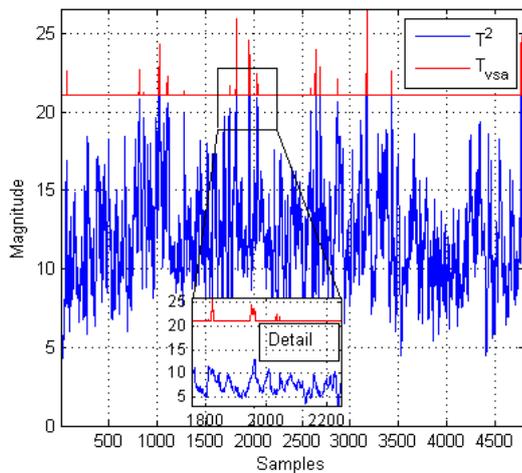


Fig. 6. T^2 and Q charts without faults.

Similar results were presented by [2], where a comparison of using different methods for data-driven fault diagnosis as the PCA, ICA, and PLS among others, were applied to the TE process, in comparison with their Table 11, almost all possible faults can be detected with a fault detection rate of 100% in this work. Although, IDV3 and IDV16 cannot be detected correctly with the T_{vsa} threshold used here.

The PCA helps in reducing the dimensionality and the correlation between variables, and it could be necessary less instrumentation, but to obtain the T^2 used to detect the faults, all the variables were needed for its computing; thus, a variable selection can be used to have a faster algorithm. The variable selection used in this paper is shown in next section.

B. PCA FAULT ISOLATION

For fault isolation, the PCA model helps to describe the normal operation condition in a process, and then by comparing the actual projections against the PCA model, the faults can be isolated. The projections give the data reconstruction removing certain variables, in the case that the variable removed is faulty, and a reconstruction of the normal operation behavior is obtained.

Some works have been made for FDI using PCA, also there are classical PCA modifications to improve the faults isolation as Interval-PCA [24], Kernel PCA [26], radii PCA [27] or dynamic PCA [28]. In [21] is shown some indices to detect and to isolate faults in their Table 4, the best results were obtained using the SPE and ϕ indices; in a similar case, the work made in [29] where some PCA methods for fault isolation were applied in a paper machine, where the best results were obtained using the residual subspace to compute the SPE, these results are shown in its Table 8. Thus, here is described the FDI using the residual subspace obtained with the PCA. As mentioned, the Q chart named SPE and described by (10), the residual subspace \tilde{X} (17) is obtained using the loading matrix P (3) as follows:

$$\tilde{X} = \bar{X}(I - PP^T). \quad (17)$$

Now by defining $C = PP^T$, a projection matrix $\Gamma^{(i)}$ is computed using (18)

$$\Gamma^{(i)} = I + \frac{\zeta_i \zeta_i^T}{1 - \zeta_i^T C \zeta_i} (C - I), \quad (18)$$

where $\zeta_i \in R^{41}$ is a vector with all elements equal to zero except the i -th which is 1.

With the projection matrix (18), the variables reconstruction is obtained using (19), where the i -th variable removed [28].

$$\hat{X}^{(i)} = \bar{X} \Gamma^{(i)}. \quad (19)$$

To isolate faults is possible to use the reconstructions from the projections made with the residual subspace, this is obtained with the residual projection matrix $\Gamma^{(i)}$ defined by (20)

$$F^{(i)} = (I - C)\Gamma^{(i)}, \quad (20)$$

the reconstructed residuals are computed using (21)

$$\tilde{X}^{(i)} = \bar{X} F^{(i)}. \quad (21)$$

The fault isolation can be achieved by removing some variables to perform the reconstructed residuals $\tilde{X}^{(i)}$, in [29] was used an incidence matrix which is orthonormal for the null space of $(I-C)$, a similar method was shown in [25], they used a matrix to denote the fault signature. In this work, the fault signature was proposed selecting the reconstructed residuals that respond to the specific fault added to the system.

In the TE process model, the faults are applied by adding a 1 into an input vector of length 20 to denote which fault will be simulated and the instant time of the fault application. The

faults magnitude cannot be controlled, but a reconstructed vector $\tilde{x}_{(i)}$ can be decomposed as (22)

$$\tilde{x}_i(k) = \tilde{x}(k) - \zeta_i F_i(k), \quad (22)$$

where $F_i(k)$ is the fault magnitude [25], considering that the i -th variable is faulty, and k represents the k -th sample. In this case, the estimated fault magnitude can be estimated using certain detection index, in this case $SPE = \tilde{P}\tilde{P}^T = I - C$. Hence, $F_i(k)$ is computing by (23)

$$F_i(k) = (\zeta_i^T (I - C) \zeta_i)^{-1} \zeta_i^T (I - C) \tilde{x}(k), \quad (23)$$

here, k represents the sample time.

The result was similar to the shown by [17], in this case was also taken into consideration the isolation time required to isolate each fault using PCA. In Table 1 are shown the detection and isolation times presented as the number of samples required. Just some works show the detection time, here is compared the results obtained in [30].

III. RESULTS AND DISCUSSION

The FDR obtained was similar to the results reported in the literature, but in this work is shown the samples needed to detect and isolate a fault because a fast detection is important for a safety process [31], one detected a fault is necessary to isolate it to know what direction to take in order to fix the problem. In Table 1 are show the samples needed for fault detection after the fault simulation, for some faults the detection is fast, but for the incipient faults the detection is achieved some hours after the fault application, these cases need to be improved because the fault can evolve affecting the process and the operators near the process installation [32], also it is possible to see that the isolation occurs soon after the fault detection. For comparison the detection times obtained from [30] are shown in the second column.

In literature is mentioned that using the SPE (Q chart) the detection has better results [3]. Another option is to improve results using soft computing to adapt the PCA. The results obtained in [33] show the FDI for faults IDV1, IDV2, IDV4, and IDV5 in the TE process with an adaptive threshold. In our opinion a mixing detection using ICA could need fewer samples until fault detection, keeping in mind that the sample time is three minutes.

Most papers use the fault detection rate (FDR) to evaluate the detection performance of the methods proposed, here in Table 2 is shown a comparison of the current application against the result shown in [24] where sparse PCA was used for fault diagnosis, the work shown in [30] where a dynamic PCA was implemented and the results of [34] where a neural network was used.

IV. CONCLUSIONS

The PCA is a common technique used for dimension reduction, it helps to avoid redundancy in the information, the main drawback is to guarantee a Gaussian distribution in the data collected, the TE model adds Gaussian noise to all the variables, but the faults could have a non-gaussian behavior, thus the ICA could need fewer samples to detect a fault. In this work were shown the detection times for the 20 faults

that can be simulated in the TE process, particularly the fault IDV9 was detected and isolated. The fault isolation was few samples after the fault detection, and the efforts will be in reducing the detection times because in four faults the detection times were bigger than the reported in the literature. Also was presented the fault detection rates that were pretty similar to those shown in similar papers. As future work, it is proposed to implement an ICA model to obtain the Q chart for the FDI. Using ICA is possible to reduce the detection times and to isolate the incipient faults IDV3 and IDV16.

TABLE I. NUMBER OF SAMPLES NEEDED FOR FDI IN THE TE PROCESS

Fault	Detection time [30]	Detection time	Isolation time
IDV1	12	33	34
IDV2	45	17	24
IDV3	Not detected	Not detected	Not isolated
IDV4	6	4	7
IDV5	6	6	7
IDV6	6	2	11
IDV7	6	2	3
IDV8	57	425	426
IDV9	Not detected	604	605
IDV10	180	659	663
IDV11	18	150	151
IDV12	63	124	126
IDV13	114	503	505
IDV14	9	94	95
IDV15	Not detected	2	4
IDV16	84	Not detected	Not isolated
IDV17	48	98	101
IDV18	27	351	355
IDV19	132	121	130
IDV20	108	197	201

TABLE II. FAULT DETECTION RATES IN THE TE PROCESS

Fault	FDR [24]	FDR [30]	FDR [34]	FDR this work
IDV1	100	99.50	97.70	98.50
IDV2	98	98.13	97.20	99.50
IDV3	7	0	0	0
IDV4	100	99.75	99.60	100
IDV5	100	21.63	92.80	100
IDV6	100	99.75	99.50	100
IDV7	100	99.75	99.50	100
IDV8	98	96.75	62.30	76.50
IDV9	6	0	0	16.75
IDV10	91	32.38	81.80	65.80

IDV11	82	86.75	52.50	72.50
IDV12	100	97.38	55.30	84.75
IDV13	95	95.25	50.40	75.60
IDV14	100	99.63	69.50	82.60
IDV15	10	0	0	100
IDV16	94	28.75	77.90	0
IDV17	97	95.63	71.40	94.55
IDV18	91	98.88	84.30	82.60
IDV19	95	8.38	97.20	77.85
IDV20	91	48.63	73.40	75.25

REFERENCES

- [1] S. Krishnannair and C. Aldrich, "Fault detection in the Tennessee Eastman benchmark process with nonlinear singular spectrum analysis," *IFAC PaperOnLine*, vol. 50, no. 1, 2017, pp. 8005–8010.
- [2] S. Yin, S. Ding, A. Haghani, H. Hao and P. Zhang, "A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process," *Journal of Process Control*, vol. 22, 2012, pp. 1567–1581.
- [3] Y. Du and D. Du, "Fault detection and diagnosis using empirical mode decomposition based principal component analysis," *Computers and Chemical Engineering*, vol. 115, 2018, pp. 1–21.
- [4] J. Chen, W. Zhang, and H.V. Poor, "An FDR-oriented approach to multiple sequential fault detection and isolation," in 2017 55th Annual Allerton Conference on Communication, Control and Computing, Monticello, 2017, pp. 112–125.
- [5] K. Khakipour, A. Safavi and P. Setoodeh, "Bearing fault diagnosis with morphological gradient wavelet," *Journal of the Franklin Institute*, vol. 354, 2017, pp. 2465–2476.
- [6] R. Gopinath, C. Santhosh-Kumar, K. Ramachandran, V. Upendranath and P. Sai-Kiran, "Intelligent fault diagnosis of synchronous generators," *Expert Systems with Applications*, vol. 45, 2016, pp. 142–149.
- [7] F. Zhou, J. Park and Y. Liu, "Differential feature based hierarchical PCA fault detection method for dynamic fault," *Neurocomputing*, vol. 202, 2016, pp. 27–35.
- [8] C. Lau, K. Ghosh, M. Hussain and C.C. Hassan, "Fault diagnosis of Tennessee Eastman process with multi-scale PCA and ANFIS," *Chemometrics and Intelligent Laboratory Systems*, vol. 120, 2013, pp. 1–14.
- [9] M.Z. Sheriff, M. Mansouri, M.N. Karim, H. Nounou and M. Nounou, "Fault detection using multi-scale PCA-based moving window GLRT," *Journal of Process Control*, vol. 54, 2017, pp. 47–64.
- [10] J. Downs and E. Vogel, "A plant-wide industrial process control problem," *Computers and Chemical Engineering*, vol. 17, no. 3, 1993, pp. 245–255.
- [11] H. Chen, P. Tinö and X. Yao, "Cognitive fault diagnosis in Tennessee Eastman process using learning in the model space," *Computer and Chemical Engineering*, vol. 67, 2014, pp. 33–42.
- [12] D.V. Ramana and S. Baskar, "Incipient fault detection of the inverter fed induction motor drive," *International Journal of Power Electronics and Drive Systems*, vol. 8, no. 2, 2017, pp. 722–729.
- [13] A. Bathelt, N. Ricker and M. Jelali, "Revision on the Tennessee Eastman process model," *IFAC Papers-Online*, vol. 48, no. 8, 2015, pp. 309–314.
- [14] X. Gao and J. Hou, "An improved SVM integrated GS-PCA fault diagnosis approach of the Tennessee Eastman process," *Neurocomputing*, vol. 174, 2016, pp. 906–911.
- [15] M.Z. Sheriff, N. Basha, M.N. Karim, H. Nounou and M. Nounou, "Fault detection of single and interval valued data using statistical process monitoring techniques," in book: *Fault Detection, Diagnosis and Prognosis*, University of Castile, 2020, pp. 1–21.
- [16] A. Alkaya and I. Eker, "Variance sensitive adaptive threshold-based PCA method for fault detection with experimental application," *ISA Transactions*, vol. 50, 2011, pp. 287–302.
- [17] K. Liu, Z. Fei, B. Yue, J. Liang and H. Lin, "Adaptive sparse principal component analysis for enhanced process monitoring and fault isolation," *Chemometrics and Intelligent Laboratory Systems*, vol. 146, 2015, pp. 426–436.
- [18] V.K. Kandula. *Fault detection in process control plants using principal component analysis*. Master's thesis, Louisiana State University and Agricultural and Mechanical College, 2011.
- [19] J.L. Devore. *Probability and statistics for engineering and the science*, 9nd ed., Cengage Learning, Boston, USA, 2014, pp. 354–356.
- [20] N. Ayech, C. Chackour and M.F. Harkat, "New adaptive moving window PCA for process monitoring," *IFAC PapersOnLine*, vol. 45, no. 20, 2012, pp. 606–611.
- [21] B. Mnassari, E. El-Adel and M. Ouladsine, "Generalization and analysis of sufficient conditions for PCA-based fault detectability and isolability," *Annual Reviews in Control*, vol. 37, 2013, pp. 154–162.
- [22] A. Casavola and G. Gagliardi, "Fault detection and isolation of electrical induction motors via LPV fault observers: a case study," *International Journal of Robust and Nonlinear Control*, vol. 25, 2015, pp. 627–648.
- [23] K. Salahshoor and F. Kiasi, "On-line process monitoring based on wavelet-ICA methodology," *IFAC Proceedings Volumes*, vol. 41, no. 2, 2008, pp. 7413–7420.
- [24] L. Lou, S. Bao and C. Tong, "Sparse robust principal component analysis with applications to fault detection and diagnosis," *Industrial & Engineering Chemistry Research*, vol. 58, 2019, pp. 1300–1309.
- [25] T. Ait-Izem, M. Harkat, M. Djeghaba and F. Kratz, "On the application of interval PCA to process monitoring: A robust strategy for sensor FDI with new efficient control statistics," *Journal of Process Control*, vol. 63, 2018, pp. 29–46.
- [26] C. Chakour, A. Benyounes and M. Boudiaf, "Diagnosis of uncertain nonlinear systems using interval kernel principal component analysis: Application to a weather station," *ISA Transactions*, vol. 83, 2018, pp. 126–141.
- [27] M. Mansouri, M. Harkat and M.N.H. Nounou, "Midpoint-radii principal component analysis-based EWMA and application to air quality monitoring network," *Chemometrics and Intelligent Laboratory Systems*, vol. 175, 2018, pp. 55–64.
- [28] E. Vanhatalo, M. Kulahci and B. Bergquist, "On the structure of dynamical principal component analysis used in statistical process monitoring," *Chemometrics and Intelligent Laboratory Systems*, vol. 167, 2017, pp. 1–11.
- [29] H. Cheng, M. Nikus and S. Jämsä, "Evaluation of PCA methods with improved fault isolation capabilities on a paper machine simulator," *Chemometrics and Intelligent Laboratory Systems*, vol. 92, 2008, pp. 186–199.
- [30] R.T. Samuel and Y. Cao, "Dynamic latent variable modeling and fault detection of Tennessee Eastman challenge process," in 2016 IEEE International Conference on Industrial Engineering (ICIT), Taipei, 2016, pp. 842–847.
- [31] Y.H. He, Y. Zhao, X. Hu, X.N. Yan, Q.X. Zhu and Y. Xu, "Fault diagnosis using novel AdaBoost based discriminant locality preserving projections with resamples," *Engineering Applications of Artificial Intelligence*, vol. 91, 2020, pp. 103631.
- [32] F. Serdio, E. Lughofer, K. Pichler, T. Buchegger and H. Efenic, "Residual-based fault detection using soft computing techniques for condition monitoring at rolling mills," *Information Science*, vol. 259, 2014, pp. 304–320.
- [33] X. Gao and J. Hou, "An improved SVM integrated GS-PCA fault diagnosis approach of Tennessee Eastman process," *Neurocomputing*, vol. 174, 2016, pp. 906–911.
- [34] D. Xie and L. Bai, "A hierarchical deep neural network for fault diagnosis on Tennessee Eastman process," in 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), Miami, 2015, pp. 745–748.
- [35] I. Prasojo, A. Maseleno, O. Tanane and N. Shahu, "Design of automatic watering system based on Arduino," *Journal of Robotics and Control*, vol. 1, no. 2, 2020, pp. 55–58.