

Combination of analytical and statistical models for dynamic systems fault diagnosis

Anibal Bregon¹, Diego Garcia-Alvarez², Belarmino Pulido¹, and Maria Jesus Fuente²

¹ *Department of Computer Science, University of Valladolid, Valladolid, Spain*
anibal@infor.uva.es, belar@infor.uva.es

² *Department of Systems Engineering and Automatic Control, University of Valladolid, Valladolid, Spain*
dieggar@cta.uva.es, maria@autom.uva.es

ABSTRACT

Complex industrial and aerospace systems require efficient monitoring and fault detection schemes to ease prognosis and health monitoring tasks.

In this work we rely upon the model-based approach to perform robust fault detection and isolation using analytical and statistical models. We have combined Principal Component Analysis (PCA) together with Possible Conflicts (PCs), to improve the overall diagnosis process for complex system. Our proposal uses residuals computed using PCs as the input for the PCA tool. The PCA tool is able to accurately determine significant deviations in the residuals, that will be identified as faults.

The integration of both techniques provides more robust results for fault detection, while avoiding false alarms in PCAs due to changes in operation modes. Moreover, it provides the PCA approach with the necessary mechanisms to perform fault isolation.

This approach has been tested on a laboratory plant with real data, obtaining promising results.

1. INTRODUCTION

Complex industrial and aerospace systems require efficient monitoring and fault detection schemes to ease prognosis and health monitoring tasks. Nevertheless, accurate and fast real-time monitoring of such system can be compromised due to the complexity of the system and the size of the measurement space. In fact, monitoring and fault detection of complex systems usually requires the integration of several techniques coming from different research fields such as knowledge-based, case-based, model-based reasoning, or machine-learning. In the remainder of this work we will focus on model-based reasoning using statistical and analytical-based on first principles models, and how to combine the best of both approaches to improve fault detection robustness.

The first technique is Principal Component Analysis (PCA) (Jackson, 1991). PCA is a well-know multivari-

This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ate statistical process control tool used in many industrial processes. PCA is able to transform a complex multivariate space into a new space with the minimum number of variables required to explain the process variation (known as latent variables). This property makes PCA a suitable tool for accurate monitoring of complex systems (Kourti & MacGregor, 1996). Moreover, as PCA is able to easily and quickly detect process variations, it has also been used within different robust fault detection schemes with successful results (Kourti & MacGregor, 1996).

Nevertheless, PCA has important limitations when dealing with continuous processes that go through different operating modes. Changes in operating modes can be caused by changes in product specifications, variation in the input variables, or modification in reference values or set-points of the system. In all these situations, the covariance structure of the process variables will be changed, leading to wrong fault detections. Another important flaw underlying this approach is that PCA provides little support for fault isolation (Gertler, Li, Huang, & McAvoy, 1999). PCA is able to detect faults as well as the set of variables involved in such fault, but this cannot be interpreted as an isolation or diagnosis stage.

On the other hand, diagnosis approaches for online fault diagnosis based on analytical models (Blanke, Kinnaert, Lunze, & Staroswiecki, 2006; Gertler, 1998) require quick and robust detection methods to determine significant deviations between observed and expected behavior. These deviations are computed using residuals, which are related to analytical redundancy derived from the system model. The structure of these residuals can be computed off-line. However, the current value of the residual is computed on-line. Whenever the value of a residual exceeds a given threshold, a fault detection is performed and the set of constraints used to derive the analytical redundancy expression is considered to be non-consistent with observations. After this process, the fault isolation is straight forward, and a reduced set of faulty candidates can be easily computed.

Residuals can be computed using different methods, such as parity-equations, state-observers, or parameter estimations. These approaches have been demonstrated to be equivalent for linear systems (Gertler, 1998). In the remainder of this work we will focus on Analytical Re-

dundancy Relations (ARRs) obtained through structural analysis, that can be used for residual generation. More precisely, our work will use Possible Conflicts (Pulido & Alonso-González, 2004), PCs, which are equivalent to ARR, and have been developed within the Artificial Intelligence community approach to Model-based Diagnosis. Possible Conflicts are able to decompose the system into the smaller number of minimal structural overdetermined sets of equations required for fault diagnosis.

In this work, we combine PCA together with Possible Conflicts, to improve the overall diagnosis process for complex systems. Our proposal uses residuals computed using PCs as the input for the PCA tool. The PCA tool will be able to accurately determine significant deviations in the residuals, that will be identified as faults. Changes in operating modes do not cause deviations in the residuals, hence we will avoid problems related with false alarms by the PCA tool when used alone. Then, after the fault detection, contribution analysis is used to determine the variables, i.e. the residuals of PCs, responsible for such deviations. Deviated residuals will be used to compute the set of faulty candidates.

The proposed combination of techniques has been tested on a laboratory plant with real data obtaining satisfactory results.

The rest of the paper is organized as follows. Sections 2 and 3 briefly introduce the Principal Component Analysis and the Possible Conflicts approaches. Then, Section 4 presents the proposed integration scheme, and Section 5 describes the experimental results obtained for the laboratory plant, a controlled two tank system. Finally, Section 6 presents the discussion and conclusions of this work.

2. PRINCIPAL COMPONENT ANALYSIS

Principal Component Analysis (PCA) is probably one of the most popular multivariate statistical techniques. PCA has been used in several fields and it has been exploited as a useful tool for fault detection in areas like Multivariate Statistical Process Control (MSPCA) (Kourti & MacGregor, 1996) and Fault Detection and Isolation (FDI) (Venkatasubramanian, Rengaswamy, Kavuri, & Yin, 2003).

PCA can be seen as an improvement over conventional univariate statistical process control (SPC). These techniques can only operate over one variable, what makes this solution not very suitable for complex industrial processes with many measured variables. The PCA approach finds linear combinations of variables that describe major trends in a data set creating a data-based model (PCA model). This model is used to monitor the process state through two statistics charts, i.e. the global behavior can be summarized in two charts.

Mathematically, PCA is a linear vector space transformation. It is performed to transform a multivariable space into a subspace which preserves maximum variance of the original space in a minimum number of dimensions capable to explain trends of the processes. These transformations are possible because the measured variables of a controlled process are highly correlated to each other.

Process data collected from the plant under nominal situation can be arranged into a matrix $\mathbf{X} \in \mathbb{R}^{K \times J}$ containing K samples of J process variables. If all vari-

ables have the same level of importance, columns of matrix \mathbf{X} can be normalized to mean 0 and variance 1 with the scale vectors $\bar{\mathbf{x}} \in \mathbb{R}^{1 \times J}$ and $\mathbf{s} \in \mathbb{R}^{1 \times J}$ (the mean and variance vectors respectively). Using this matrix the principal components can be computed through the covariance matrix $\mathbf{R} \in \mathbb{R}^{J \times J}$ (Wold, 1987):

$$\mathbf{R} = \frac{1}{K-1} \mathbf{X}^T \mathbf{X} \quad (1)$$

performing the singular value decomposition (SVD) over \mathbf{R} (Chiang, Russell, & Braatz, 2000):

$$\mathbf{R} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T, \quad (2)$$

where $\mathbf{\Lambda} \in \mathbb{R}^{J \times J}$ and $\mathbf{V} \in \mathbb{R}^{J \times J}$ are the results of SVD decomposition. $\mathbf{\Lambda}$ is a diagonal matrix that contains the eigenvalues of \mathbf{R} in its diagonal $\lambda_1, \lambda_2, \dots, \lambda_J$, sorted in decreasing order. The columns of matrix \mathbf{V} are the eigenvectors of \mathbf{R} . The transformation matrix $\mathbf{P}_{1:A} \in \mathbb{R}^{J \times A}$ is arranged selecting A eigenvectors or columns of \mathbf{V} corresponding to the A greatest eigenvalues. The matrix $\mathbf{P}_{1:A}$ transforms the space of the measured correlated variables into the reduced dimension space of uncorrelated variables. This matrix is also called the *loadings* matrix. The space transformation can be expressed as follows:

$$\mathbf{T} = \mathbf{X} \mathbf{P}_{1:A} \quad (3)$$

$\mathbf{T} \in \mathbb{R}^{K \times A}$ is the *scores* matrix, this matrix is made up of the original process variables transformed into the reduced dimension space. Every new variable of this space is called the a -th score t_a .

Using Eq. (3), the scores can be transformed into the original space:

$$\hat{\mathbf{X}} = \mathbf{T} \mathbf{P}_{1:A}^T \quad (4)$$

The residual matrix \mathbf{E} is calculated as:

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}} \quad (5)$$

Finally, the original data space can be calculated as:

$$\mathbf{X} = \mathbf{T} \mathbf{P}_{1:A}^T + \mathbf{E} \quad (6)$$

Equation (6) is a summary of PCA, i.e. data matrix \mathbf{X} can be expressed as the sum of two terms. The first term $\mathbf{T} \mathbf{P}_{1:A}^T$ is the PCA model. This term can be considered as an estimation of \mathbf{X} . The second term \mathbf{E} represents the noise in the process, and it is rejected.

A decisive task within this approach is to choose accurately the number of principal components, A . There are different strategies to carry out this task (Jackson, 1991; Chiang et al., 2000; Weighell, Martin, & Morris, 2001). The most suitable procedure to choose the number of principal components is cross validation (Eastment & Krzanowski, 1982; Bro, Kjeldahl, & Kiers, 2008). This method is based on the selection of the components which maximize the goodness of fit and the goodness of prediction of the PCA model.

2.1 Monitoring statistic and fault detection

When the PCA model is fitted, the *on-line* monitoring using multivariate control charts based on Hotelling's T^2 and square prediction error (SPE) or Q can be performed. The monitoring task is reduced to these charts (T^2 and Q).

T^2 statistic is computed for a new measured variables vector \mathbf{x} as follows:

$$T^2 = \mathbf{x}^T \mathbf{P}_{1:A} \Lambda_A^{-1} \mathbf{P}_{1:A}^T \mathbf{x} \quad (7)$$

where Λ_A is a squared matrix arranged by the A rows and columns of Λ .

When the process is monitored using this statistic, it can be considered *normal* for a specific level of significance α if:

$$T^2 \leq T_\alpha^2 = \frac{(K^2 - 1)A}{K(K - A)} F_\alpha(A, K - A) \quad (8)$$

In Eq. (8), $F_\alpha(A, K - A)$ is the critical value of the Fisher-Snedecor statistical distribution with K and $K - A$ degrees of freedom and α is the level of significance. α is typically assigned to values between 90% and 95%.

As Eq. (7) Hotelling's statistic is computed using the first A larger principal components. It means that T^2 can be seen as a test for deviations in the latent variables, this is usually due to changes in the correlations structure of the original variables. Hence, T^2 statistic is triggered when the variation observed in the latent variables is larger than the variation explained by the common causes. This drawback is solved through the squared prediction error SPE statistic, also known as Q statistic. This statistic is calculated as the sum of the squares of the residuals (Jackson & Mudholkar, 1979; Jackson, 1991). The Q statistic is a measurement of *goodness of fit* of the sample to the model and it is directly associated with noise:

$$Q = \mathbf{r}^T \mathbf{r} \quad (9)$$

with:

$$\mathbf{r} = (\mathbf{I} - \mathbf{P}_{1:A} \mathbf{P}_{1:A}^T) \mathbf{x}$$

The threshold for this statistic is computed as follows:

$$Q_\alpha = \theta_1 \left[\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}} \quad (10)$$

with:

$$\theta_i = \sum_{j=a+1}^m \lambda_j^i \quad h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2}$$

In Eq. (10), c_α is the value of the normal statistical distribution, with α level of significance.

Unusual events that produce changes in the covariance structure of original variables are detected by a high value of Q .

2.2 Contribution analysis for fault isolation

Even though PCA has been successfully used for monitoring and fault detection as shown in the previous section, the classical PCA approach provides little information for fault isolation. As a first approximation for fault isolation, it has been proposed the contribution analysis (Kourti & MacGregor, 1996). The contribution analysis automatically computes the influence of each one of the components of the PCA within the behavior of the Q and T^2 statistics. One of the most used approaches for contribution analysis is to study the normalized error of the variables.

Regarding the Q statistic, the contribution analysis is triggered when an observation $\mathbf{x} \in \mathbb{R}^{1 \times J}$ makes the Q statistic to exceed its threshold. The normalized error is computed as follows:

$$\mathbf{e} = \frac{\mathbf{x} - \bar{\mathbf{x}}}{\mathbf{s}} \quad (11)$$

The contribution of variables is considered as normal when its value, assuming normal distribution, has values out of $\sigma \geq 3$, i.e. approximately 99% of the samples should fall into this interval. Variables in \mathbf{e} with values out of this interval are interpreted as the variables responsible for the deviation in the Q statistic. For practical reasons, the normalized error vector $\mathbf{e} \in \mathbb{R}^{1 \times J}$ is usually represented on a common bar plot.

The bar plot of the normalized error of the variables cannot be used to diagnose the fault if it is too small and the variables are highly correlated. To deal with such drawback the bar plot of normalized scores is computed. In this case every normalized score is computed as follows:

$$\bar{t}_a = \frac{t_a}{\lambda_a} \quad (12)$$

The bar plot of normalized scores is made up of the normalized scores vector $\bar{\mathbf{t}} = [\bar{t}_1, \bar{t}_2, \dots, \bar{t}_A]$. When the T^2 statistics value is larger than its threshold, the normalized scores with high value (when are larger than ± 3 as it happens with the Q statistics) are considered responsible for the deviation in the T^2 statistic. Then, variables contribution to individual scores are computed to identify the variables responsible of the deviation. For practical reasons, variables that contribute to the computation of every score $\mathbf{x}_t = [x_{t, a1}, x_{t, a2}, \dots, x_{t, aJ}]$ are represented on a common bar plot. Each one of these contributions is computed as the product of the value of the original variable x_j with the corresponding loading for this score p_{aj} :

$$x_{t, aj} = p_{aj} x_j \quad (13)$$

3. POSSIBLE CONFLICTS DIAGNOSIS APPROACH

Possible Conflicts, or PCs (Pulido, Alonso, & Acebes, 2001; Pulido & Alonso-González, 2004), is an off-line dependency compilation technique from the Artificial Intelligence community (also known as the DX community). PCs are minimal subsets of equations with enough redundancy to perform fault diagnosis. The main idea underlying Possible Conflicts is that all those subsystems capable to become a conflict can be identified off-line. In

FDI terms, a PC represents the structure of an ARR that can be used for fault detection and isolation.

Computation of PCs is done off-line by using an abstract representation of the system: an hypergraph (Pulido et al., 2001) or a bond graph model of the system (Bregon, Pulido, Biswas, & Koutsoukos, 2009). Classically, Possible Conflicts are computed by a two-step process:

- Computation of the Minimal Structural Overdetermined sets of constraints. Each one of these sets, known as Minimal Evaluation Chains, MECs, represents a necessary condition for a conflict to exist. Each constraint within a MEC contains one or more variables. When a variable inside a constraint can be solved assuming the rest of the variables are known, this is called an interpretation, i.e. a feasible causal assignment, and this leads to the second step.
- Search for all the causally consistent interpretations for each constraint in a MEC, which is called the Minimal Evaluation Model, MEM. Each MEM represents a globally consistent causal assignment within a MEC and can be used to estimate the behavior of a part of the whole system.

Moreover, since conflicts only arise when models are evaluated using the set of available observations, the set of constraints within a MEM is called a *Possible Conflict*.

Minimal Evaluation Models can be used to perform fault detection by looking for discrepancies between estimated variables and the observed ones. If there is a discrepancy, the Possible Conflict would be responsible for such a discrepancy and it will be confirmed as a real conflict. Then, diagnosis candidates (faulty components) are obtained following Reiter's theory (Reiter, 1987).

It has been proven that PCs computation is equivalent, under certain circumstances, to the on-line conflict computation in the General Diagnostic Engine, GDE, and to the off-line generation of ARRs in the Control Theory approach to diagnosis (Pulido & Alonso-González, 2004; Armengol et al., 2009).

A detailed description of the Consistency-based Diagnosis approach using PCs can be found in (Pulido et al., 2001).

4. ROBUST FAULT DIAGNOSIS BY INTEGRATION OF PCA AND PCS

The classical PCA has been successfully used as a tool for robust fault detection in complex industrial plants, but, as described before, it faces problems when dealing with the changing nature of industrial processes. Changes in the operation conditions trigger a large number of false alarms (or missed detections if these false alarms want to be avoided (Zumoffen & Basualdo, 2007)). Currently, there exist several modifications and improvements to PCA proposed to deal with such drawback, like the multi-scale PCA (MSPCA) (Misra, Yue, Qin, & Ling, 2002), the adaptive PCA (APCA) (Zumoffen & Basualdo, 2007), the recursive PCA (Li, Yue, Valle-Cervantes, & Qin, 2000), the exponentially weighted PCA (EWPCA) (Lane, Martin, Morris, & Gower, 2003), and the dynamic PCA (Ku, Storer, & Georgakis, 1995). All these solutions can be classified into the three following categories ((Hwang & Han, 1999) (Tien, Lim, & Jun, 2004)):

- a) Build a PCA model for each operation mode.
- b) Update the model to reflect the changes in the operation modes.
- c) Develop a conventional PCA model to account for all such changes.

The solution proposed in this paper do not fall into any of these categories. In this case, a classical PCA approach will be applied over the residuals produced by a model-based fault detection and diagnosis (FDI) technique (the Possible Conflicts) instead of using the original process variables. Possible Conflicts are able to provide residuals only sensitive to a subset of faults and not sensitive to changes in operation conditions. Our integration approach exploits this property to compute PCA with similar properties.

Regarding fault diagnosis, the classical PCA do not provide the necessary mechanisms to isolate faulty candidates when a fault is detected in the system. Recently, several approaches have been proposed to overcome such problem. (Gertler et al., 1999) have proved the equivalence between PCA and parity relations, and then, have used such analogy to design structured partial PCA models with the same isolability properties than the parity relations. Using such equivalence, (Gertler et al., 1999) and (Huang, Gertler, & McAvoy, 2000) decompose the original PCA model into "smaller" structured PCA models that guarantee the disturbance decoupling for the set of faults considered. Our proposal in this paper is related with the approach proposed in (Gertler et al., 1999; Huang et al., 2000), but instead of designing off-line a set of partial structured PCA, we just design a PCA model using the residuals of the PCs, which can be considered as a special case of ARRs. Then, a contribution analysis (using contribution plots) is done automatically to obtain the residuals responsible for the deviation in the PCA model. Using the theoretical fault signature matrix provided by the PCs and the activated residuals we can isolate the fault that has occurred in the system.

Next section describes in detail our proposal to integrate PCs and PCA models.

4.1 The Integration Proposal

Figure 1 illustrates the scheme of the integration proposal. The system model is decomposed into minimal structural overdetermined subsystems by computing the set of Possible Conflicts. Then, the residuals computed by the Possible Conflicts for a small subset of training data is introduced to the **PCA model computation** block. In this block the PCA model is computed and the upper limits for the statistics are fitted. This process is carried out off-line. Then, on-line, the **residual computation** blocks (one for each PC) compute the residuals, $R_1 \dots R_i$, as the linear difference between the estimations provided by the PCs, \hat{y}_{pc_x} , and the measurements, y . These residuals are introduced into the **PCA** block that provides the fault detection results based on the computation of the T^2 and Q statistics. In our proposal for fault detection we have considered that a fault has occurred in the system whether the T^2 or the Q statistics values are larger than the detection threshold during an empirically determined number of consecutive samples.

When the **PCA** block deviates from the nominal situation, the **contributions computation** block determines

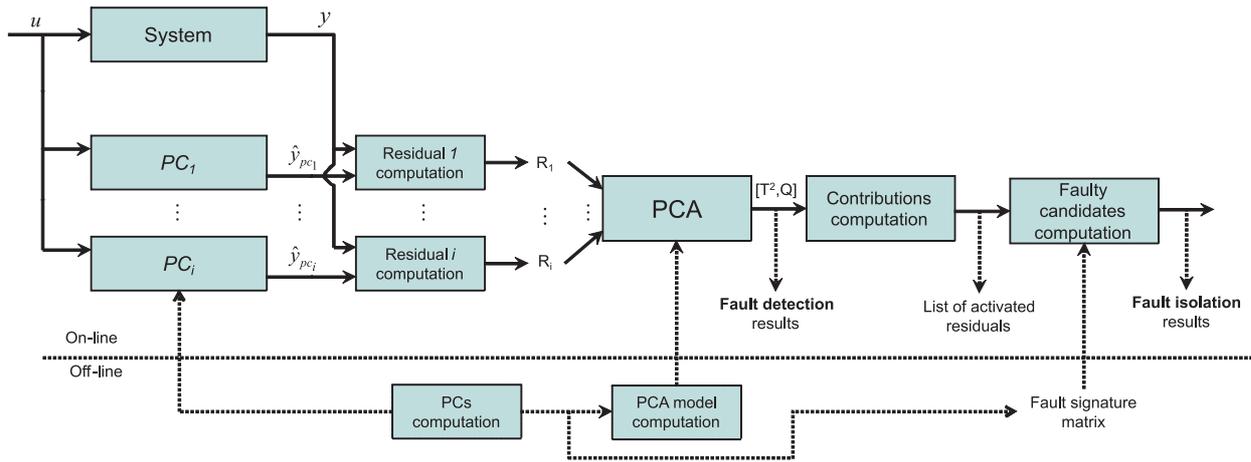


Figure 1: Integration proposal scheme.

the variables, i.e. the residuals, responsible for such deviation. We used contribution analysis techniques, as described in Section 2, to implement the **contributions computation** block. As a final step, the set of faulty candidates is computed by a minimal hitting set procedure of the residuals responsible for the fault using the fault signature matrix computed by the Possible Conflicts.

It is important to point out that the computation effort required for this approach is quadratic in time for on-line fault detection. Computing linear distance for the residuals is linear in time, and then computation of both T^2 and Q statistics entails a process of matrices multiplication (as shown in Section 2) that is polynomial. Regarding on-line fault isolation, computational complexity is determined by the minimal hitting set algorithm, that it is cubic when an incremental minimal hitting set algorithm is considered (Dressler & Struss, 1996).

Next section shows the experiments we conducted on real data for a laboratory plant to prove the validity of the proposed diagnosis scheme.

5. RESULTS ON THE CASE STUDY

The laboratory plant we used to test the integration scheme proposed in the previous section is a controlled two-tank system. We run several experiments using real data for different magnitudes and time fault occurrence for sensor faults.

5.1 Description of the Laboratory Plant

The laboratory plant (shown in Figure 2) is made up of two water tanks, T1 and T2, both with the same transversal area. Two PID controllers try to keep the level of the tanks close to a reference level acting over two pumps, P1 and P2. The pumps provide the input flow q_1 and q_2 to the tanks T1 and T2, respectively. Both tanks are connected through a valve, q_{12} . The level of the tanks is measured through two level sensors, h_1 (named LC1 in Figure 2) that measures the level of tank T1, and h_2 (named LC2 in Figure 2) that measures the level of tank T2. The liquid is introduced into the tanks by means of two frequency controllable electronic pumps. To model the behavior of the system we have used first-principles

equations. A more detailed description of this plant can be found in (Fuente, Garcia, & Sainz, 2008).

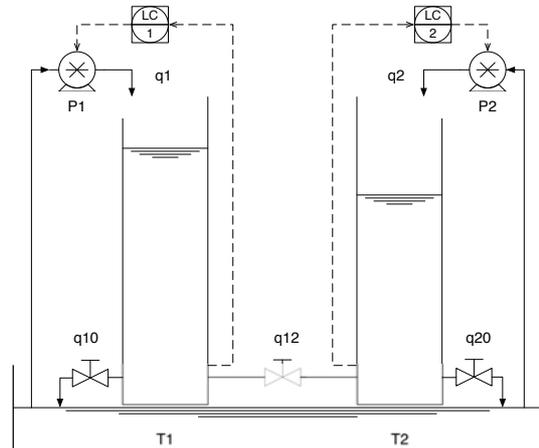


Figure 2: Two tanks system.

Two operation modes have been considered in the plant. The operation mode 1, where the level of tank T1 is set to 30% and level of tank T2 to 30%; and the operation Mode 2 where the level of tank T1 is kept to 30% and level of tank T2 is increased to 50%. Therefore, two transitory states appear during the plant operation: the transitory state between initial point and the operation mode 1; and the transitory state between operation modes 1 and 2. The change in the reference is done at time instant $t = 400$ seconds.

Monitoring techniques presented in this work have been implemented using MATLAB[©] and SIMULINK[©] tools. The data acquired from the plant (by means of a data acquisition card model PCI-DAS1002) were manipulated with these tools using the OPC[©] communication protocol. In this architecture, the data acquired by the card is presented by a OPC[©] server developed in VC++ and the monitoring applica-

tions implemented in MATLAB[©] and SIMULINK[©] run as OPC[©] clients.

5.2 PCs and PCA for the Laboratory Plant

Based on the model description shown in the previous section we have found a set of four Possible Conflicts (shown in Table 1). These Possible Conflicts are minimal w.r.t. the set of constraints in the models. In the table, the first column shows the PC identifiers, the second column illustrates the set of components involved in each PC, and finally, the third column indicates variable estimated by each Possible Conflict. PCs relation to faulty components are shown in the theoretical fault signature matrix (shown in Table 2). This matrix describes the PCs that should be triggered when a fault in a component occurs.

Table 1: PCs found for the plant: components and estimated variable for each Possible Conflict.

	Components or Support	Estimate
PC_1	$T_1, T_2, q_1, q_2, q_{12}, h_2(\text{sensor})$	h_2
PC_2	$T_1, T_2, q_1, q_2, q_{12}, h_1(\text{sensor})$	h_1
PC_3	$T_1, q_1, q_{12}, h_1(\text{sensor}), h_2(\text{sensor})$	h_1
PC_4	$T_2, q_2, q_{12}, h_1(\text{sensor}), h_2(\text{sensor})$	h_2

Table 2: PCs and their related fault modes. The set of faults considered in this plant are: faulty sensors (f_{h_1}, f_{h_2}), blockages of pipes/vales ($f_{q_1}, f_{q_2}, f_{q_{12}}$), and leakages (f_{T_1}, f_{T_2}).

	f_{q_1}	f_{q_2}	$f_{q_{12}}$	f_{T_1}	f_{T_2}	f_{h_1}	f_{h_2}
PC_1	1	1	1	1	1		1
PC_2	1	1	1	1	1	1	
PC_3	1		1	1		1	1
PC_4		1	1		1	1	1

The PCA model has been computed using the output of the PCs calculated for ten real experiments in nominal situation. These experiments were run for 1600 seconds. The PCA model has been fitted with the three principal components. Regarding the detection thresholds for the PCA model, in this work we have considered a level of significance $\alpha = 95\%$.

5.3 Results

We first considered experiments in nominal situation. Table 3 shows the mean value of the false alarms percentage we obtained for 20 experiments in nominal situation. The table compares the percentage of false alarms we obtained when using PCA alone against the integration of PCs and PCA. Looking at the table, it is clear that the number of false alarms is reduced when PCA and PCs are used together. As we previously explained, the occurrence of a fault in the system is determined by a consecutive number of fault alarms. Hence, a decrease in the number of false alarms will cause a decrease in the number of false positives in the detection. For this laboratory plant, we obtained at least one false positive in all the nominal experiments when PCA is used alone (due to start-ups and changes in the references), but no false positives in the detection when the integration proposal

is used. Figure 3 shows the output of the T^2 and Q statistics for an example in nominal situation when PCA is used alone. Figure 4 shows the same output for the same example when the integration of PCs and PCA is used. Comparing Figure 3 against Figure 4 we can clearly see that using PCA alone causes a larger number of false alarms and one false positive (during the start-up), while the combination of techniques causes a smaller number of false alarms and no false positives.

Table 3: Mean value of the false alarms percentage obtained for 20 experiments in nominal situation.

	PCA	PC + PCA
False alarms percentage:	7.75	2.01

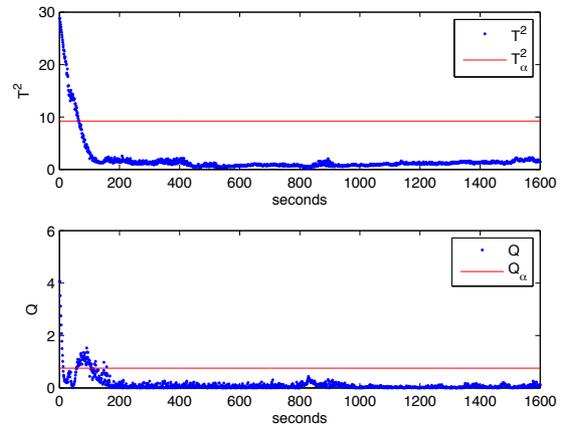


Figure 3: T^2 and Q statistics for a nominal experiment when PCA is used alone.

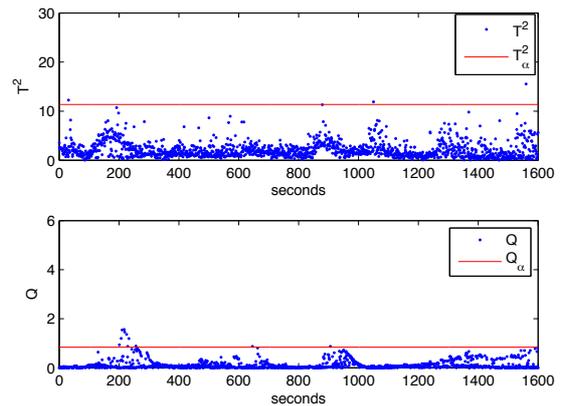


Figure 4: T^2 and Q statistics for a nominal experiment when PCA is used together with PCs.

On the other hand, Table 4 shows the results we obtained in different faulty situations. In this work we only considered sensor faults (h_1 sensor and h_2 sensor). We

considered 40% and 60% fault sizes occurring at times $t = 100$, $t = 500$, $t = 850$, or $t = 1300$ seconds. Table 4 shows the detection time using the T^2 and the Q statistics (T^2 detection time and Q detection time), as well as the time instant after fault detection when the approach is able to uniquely isolate the fault (T^2 isolation time for the T^2 statistic, and Q isolation time for the Q statistic).

Looking at the results, it is clear that the approach is able to accurately detect every fault. For all the experiments we carried out, the T^2 statistic was able to detect the faults within the first time steps after the fault occurrence (and it did it faster than the Q statistic¹). Regarding fault isolation, for all the experiments we carried out (except one, a 40% fault size at $t = 850$ in h_1 sensor), the T^2 statistic was able to uniquely isolate the faults. In this table, the results only show isolation time when the faults were uniquely isolated, but for the rest of the cases, the approach was able to isolate a small subset of faulty candidates, reducing this way the initial number of faulty candidates.

Figures 5, 6, and 7 show an example of the output of the system when a 40% fault was introduced in sensor h_1 at time $t = 500$. Figure 5 shows the output of the residuals for such fault. Using these residuals as the input for the PCA model, we obtained the output for the T^2 and Q statistics shown at Figure 6. Looking at this figure we can see that using the integration approach, the PCA did not face problems related with the start-up and it was able to quickly and accurately detect the fault. Finally, the contribution plot shown at Figure 7 illustrates the contribution analysis for the T^2 statistic. In this example, residuals for Possible Conflicts PC_2 , PC_3 , and PC_4 exceeded the ± 3 threshold, and consequently were considered as the cause of the deviation in the T^2 statistic. Looking at the theoretical fault signature matrix in Table 2 the approach determined that the faulty component was the sensor h_1 , because a fault in this sensor triggers Possible Conflicts PC_2 , PC_3 , PC_4 , and not PC_1 (a fault in sensor h_2 would have had instead triggered PC_1 and not PC_2).

6. DISCUSSION AND CONCLUSIONS

Several approaches have been proposed in the literature to reduce the dimensionality of industrial systems to ease monitoring and fault detection. Among these techniques, Principal Component Analysis has proven success. PCA is able to reduce the dimensionality of the observation space and provide robust fault detection, but it faces problems when changes in the working conditions and the operation modes occur. In this work we have proposed to integrate PCA with a model-based diagnosis approach, the Possible Conflicts, to improve the overall diagnosis process.

This approach improves the PCA-based fault detection scheme because the model-based detection system acts as a filter of changes between the different operating modes. In this case, a PCA model is built using the residual signals from the process under normal operation. When a fault occurs in the system, one or more

¹It is important to point out that the detection decision can be given by either the T^2 or the Q statistics.

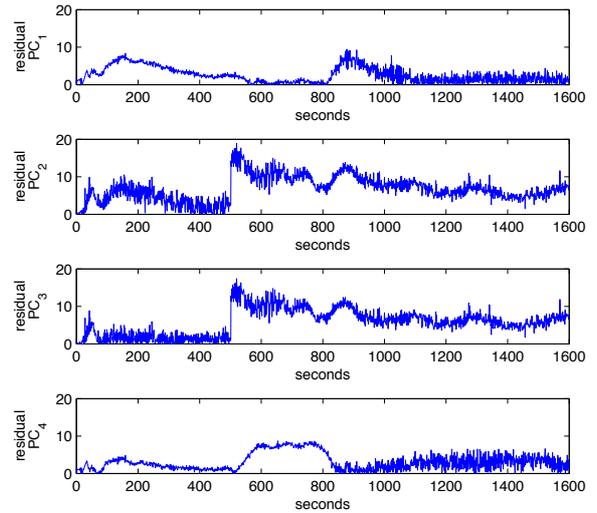


Figure 5: PCs residuals for a 40% fault in sensor h_1 at time $t = 500$.

of these residuals will suffer a deviation that will be detected by the PCA-based detection system. This detection is possible through the T^2 and Q statistics. Moreover, the integration approach also improves the PCs-based fault diagnosis scheme because the statistical analysis of the residuals provides fault detection capabilities which are robust to model uncertainties and noisy measurements.

Another improvement of the approach is related with the diagnosis tasks. The basic PCA approach provides little fault isolation capabilities. When a fault is detected, a contribution analysis is performed to identify variables responsible for the fault. But this diagnosis procedure has to be interpreted carefully because the variables are highly correlated and variables that are not involved in the fault may have a high contribution value. Using

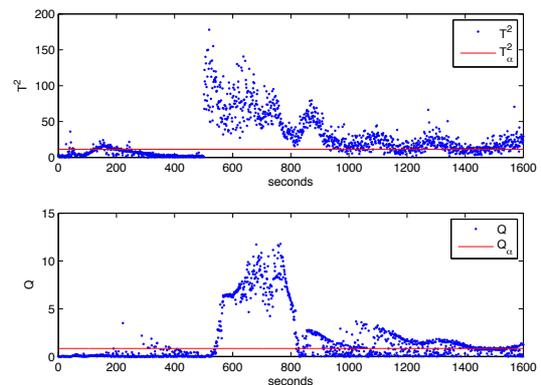
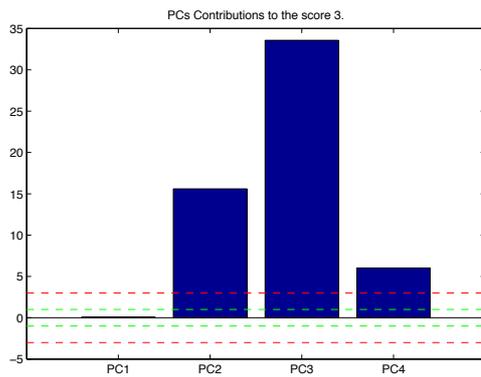


Figure 6: T^2 and Q statistics for the faulty experiment shown at Figure 5.

Table 4: Result for different faulty situations when the integration of PCA and PCs is used.

Faulty component:	h_1 sensor				h_2 sensor			
Fault instant:	$t = 100$	$t = 500$	$t = 850$	$t = 1300$	$t = 100$	$t = 500$	$t = 850$	$t = 1300$
Fault size:	40% fault size							
T^2 detection time	105	501	900	1301	101	501	901	1301
T^2 isolation time	175	571		1361	161	521	961	1661
Q detection time	164	543				579	969	
Q isolation time								
Fault size:	60% fault size							
T^2 detection time	105	501	900	1301	101	501	901	1301
T^2 isolation time	163	541	941	1361	141	521	941	1341
Q detection time	162	521	940	1365	160	524	953	1462
Q isolation time				1381		524		1462

Figure 7: Contribution plot for the T^2 statistic for the faulty experiment shown at Figure 5.

the integration of techniques proposed in this paper this problem disappears, because the contribution analysis identifies residuals (instead of variables) responsible for the deviation in the statistic. Then, a minimal hitting set algorithm is used to identify faulty components.

The integration proposal has been tested using different experiments with real data from a laboratory plant. Results obtained showed the validity of the approach. We obtained robust fault detection results, reducing the number of false alarms in nominal situation (avoiding false positives in the detection) and providing the PCA-based diagnosis approach with accurate fault isolation capabilities.

Looking at the results, main conclusion of this work is that integration of PCA with PCs increases robustness in the detection for both approaches, and allows to isolate faults with PCA.

As future work we are planning to perform a more exhaustive experimental study including different faulty situations and smaller fault magnitudes. Our guess is that the integration proposal will be able to detect smaller faults than considering PCA or PCs alone. We are also planning to test our approach also on component faults.

ACKNOWLEDGMENTS

Anibal Bregon and Belarmino Pulido's work has been partially supported by Spanish regional grant JCyL VA005B09. Diego Garcia-Alvarez and Maria Jesus

Fuente's work has been partially supported by Spanish government grant DPI2009-14410-C02-02.

NOMENCLATURE

\mathbf{X}	measured data matrix
K	number of samples
J	number of process variables
\mathbf{R}	covariance matrix
$\mathbf{\Lambda}$	matrix that contains the eigenvalues of R
\mathbf{V}	matrix that contains the eigenvectors of R
λ_j	j -th eigenvalue of R
\mathbf{P}	loadings matrix
\mathbf{T}	scores matrix
\mathbf{E}	residual matrix
T^2	Hotelling's statistic
Q	Squared prediction error statistic

REFERENCES

- Armengol, J., Bregon, A., Escobet, T., Gelso, E., Krysander, M., Nyberg, M., et al. (2009). Minimal Structurally Overdetermined sets for residual generation: A comparison of alternative approaches. In *Proceedings of the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes, SAFEPROCESS09* (p. 1480-1485). Barcelona, Spain.
- Blanke, M., Kinnaert, M., Lunze, J., & Staroswiecki, M. (2006). *Diagnosis and Fault-Tolerant Control*. Springer.
- Bregon, A., Pulido, B., Biswas, G., & Koutsoukos, X. (2009). Generating Possible Conflicts from Bond Graphs Using Temporal Causal Graphs. In *Proceedings of the 23rd European Conference on Modelling and Simulation, ECMS09* (p. 675-682). Madrid, Spain.
- Bro, R., Kjeldahl, K., & Kiers, H. (2008). Cross-validation of component models: A critical look at current methods. *Analytical and Bioanalytical Chemistry*, 390, 1241-1251.
- Chiang, L., Russell, E., & Braatz, R. (2000). *Fault Detection and Diagnosis in Industrial Systems*. Nueva York: Springer.
- Dressler, O., & Struss, P. (1996). The Consistency-based approach to automated diagnosis of devices. In G. Brewka (Ed.), *Principles of Knowledge Representation* (p. 269-314). CSLI Publications, Stanford.

- Eastment, H., & Krzanowski, W. (1982). Cross validation choice of the number of components from a principal component analysis. *Technometrics*, 24, 73-77.
- Fuente, M., Garcia, G., & Sainz, G. (2008). Fault diagnosis in a plant using Fisher discriminant analysis. *Proceeding of the 16th Mediterranean Conference on Control and Automation Congress Centre, Ajaccio, France*, 53-58.
- Gertler, J. (1998). *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, Inc.
- Gertler, J., Li, W., Huang, Y., & McAvoy, T. (1999). Isolation enhanced principal component analysis. *American Institute of Chemical Engineers (AIChE) Journal*, 45, 323-334.
- Huang, Y., Gertler, J., & McAvoy, T. (2000). Sensor and actuator fault isolation by structured partial PCA with nonlinear extensions. *Journal of Process Control*, 10(5), 459-469.
- Hwang, D., & Han, C. (1999). Real-time monitoring for a process with multiple operating modes. *Control Engineering Practice*, 7, 891-902.
- Jackson, J. (1991). *A user's guide to principal components*. Wiley.
- Jackson, J., & Mudholkar, G. (1979). Control procedures for residuals associated with principal component analysis. *Technometrics*, 21, 341-349.
- Kourti, T., & MacGregor, J. (1996). Multivariate SPC Methods for Process and Product Monitoring. *Journal of Quality Technology*, 28, 409-428.
- Ku, W., Storer, R., & Georgakakis, C. (1995, November). Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and intelligent laboratory systems*, 30, 179-196.
- Lane, S., Martin, E., Morris, A., & Gower, P. (2003). Application of exponentially weighted principal component analysis for the monitoring of a polymer film manufacturing process. *Transactions of the Institute of Measurement and Control*, 25, 17-35.
- Li, W., Yue, H., Valle-Cervantes, S., & Qin, S. (2000). Recursive PCA for adaptive process monitoring. *Journal of Process Control*, 10, 471-486.
- Misra, M., Yue, H., Qin, S., & Ling, C. (2002). Multivariable process monitoring and fault diagnosis by multi-scale PCA. *Computers & Chemical Engineering*, 26, 1281-1293.
- Pulido, B., Alonso, C., & Acebes, F. (2001). Lessons learned from diagnosing dynamic systems using possible conflicts and quantitative models. In *Engineering of Intelligent Systems. XIV Conf. IEA/AIE-2001* (Vols. LNAI, Current Topics in AI, Springer, p. 135-144). Budapest, Hungary.
- Pulido, B., & Alonso-González, C. (2004, Octubre). Possible Conflicts: a compilation technique for consistency-based diagnosis. *IEEE Trans. on Systems, Man, and Cybernetics. Part B: Cybernetics*, 34(5), 2192-2206.
- Reiter, R. (1987). A Theory of Diagnosis from First Principles. *Artificial Intelligence*, 32, 57-95.
- Tien, D., Lim, K., & Jun, L. (2004, November 2-6). Comparative study of PCA approaches in process monitoring and fault detection. *The 30th annual conference of the IEEE industrial electronics society*, 2594-2599.
- Venkatasubramanian, V., Rengaswamy, R., Kavuri, S., & Yin, K. (2003). A review of process fault detection and diagnosis. Part I, II and III: Quantitative model-based methods. *Computers & Chemical Engineering*, 27, 291-346.
- Weighell, M., Martin, E., & Morris, A. (2001). The statistical monitoring of a complex manufacturing process. *Journal of Applied Statistics*, 28, 409-425.
- Wold, S. (1987). Principal Component Analysis. *Chemometrics and intelligent laboratory systems*, 2, 37-52.
- Zumoffen, D., & Basualdo, M. (2007). From large chemical plant data to fault diagnosis integrated to decentralized fault tolerant control: pulp mill process application. *Industrial & Engineering Chemistry Research*, 47, 1201-1220.

Anibal Bregon received his B.S., M.S., and Ph.D. degrees in computer science from the University of Valladolid, Valladolid, Spain, in 2005, 2007, and 2010, respectively. Currently he is Research Assistant at the Department of Computer Science from the University of Valladolid. His main research interests are model-based reasoning for diagnosis, prognosis, and health-management.

Diego Garcia-Alvarez received his B.S. degree in computer science in 2006 and his M.S. degree in process and system engineering in 2008, both from the University of Valladolid, Valladolid, Spain. Currently he is Ph.D student at department of system engineering and automatic control from the University of Valladolid. His main research interests are monitoring and fault detection and diagnosis methods using multivariate statistical process control methods.

Belarmino Pulido received his B.S., M.S., and Ph.D. degrees in computer science from the University of Valladolid, Valladolid, Spain, in 1992, 1995, and 2001, respectively. He has been working since 1994 with the Department of Computer Science, University of Valladolid, where he is Associate Professor since 2002. His main research interests are model-based reasoning and knowledge-based reasoning, applied to diagnosis, prognosis, and health-management.

Maria Jesus de la Fuente received her Ph.D. degree in Sciences (Physics) from the University of Valladolid, Valladolid, Spain, in 1994. She has been working since 1992 with System Engineering and Automatic Control Department of the University of Valladolid, where she is Associate Professor since 1997. Her main research interests are model-based reasoning and data-driven methods, applied to diagnosis, fault tolerant control, prognosis and health-management.