



Sensitivity enhancing transformations for monitoring the process correlation structure



Tiago J. Rato, Marco S. Reis*

CIEQPF, Department of Chemical Engineering, University of Coimbra, Rua Sílvio Lima 3030-790, Coimbra, Portugal

ARTICLE INFO

Article history:

Received 31 October 2013
 Received in revised form 3 April 2014
 Accepted 4 April 2014
 Available online 1 May 2014

Keywords:

Process monitoring
 Multivariate dynamical processes
 Variables transformation
 Partial correlations
 Marginal correlations

ABSTRACT

In this article we address the general problem of monitoring the process cross- and auto-correlation structure through the incorporation of information about its internal structure in a pre-processing stage, where sensitivity enhancing transformations are applied to collected data. We have found out that the sensitivity of the monitoring statistics based on partial or marginal correlations in detecting structural changes is directly related to the nominal levels of the correlation coefficients during normal operation conditions (NOC). The highest sensitivities are obtained when the process variables involved are uncorrelated, a situation that is hardly met in practice. However, not all transformations perform equally well in producing uncorrelated transformed variables with enhanced detection sensitivities. The most successful ones are based on the incorporation of the natural relationships connecting the process variables. In this context, a set of sensitivity enhancing transformations are proposed, which are based on a network reconstruction algorithm. These new transformations make use of fine structural information of the variables connectivity and therefore are able to improve the detection capability to local changes in correlation, leading to better performances when compared to current marginal-based methods, namely those based on latent variables models, such as PCA or PLS. Moreover, a novel monitoring statistic for the transformed variables variance proved to be very useful in the detection of structural changes resulting from model mismatch. This statistic allows for the detection of multiple structural changes within the same monitoring scheme and with higher detection performances when compared to the current methods.

© 2014 Published by Elsevier Ltd.

1. Introduction

Statistical Process Control (SPC) is a pervasive industrial task with the aim to monitor the evolution of the overall process operation state. It is implemented mostly by means of control charts designed to assess whether the process is only subject to normal causes of variation, inherent to its operation, or if a special cause of variation has occurred, that urges to be detected, diagnosed and fixed or accommodated. Special causes can result from equipment failure, abnormal changes in raw materials properties or any internal or external events that affect process evolution and increase its variability, putting in risk not only product quality and process efficiency, but also the safety of people and the environment. Therefore, the rapid detection of faults and associated diagnosis of their origins are of outmost importance, and have attracted the interest of many researches and practitioners. However, even without

entering in an exhaustive enumeration of the multivariate methodologies proposed through the decades, one can easily acknowledge that the proposed control charts were essentially designed to monitor the process mean [1–5], having some limited ability to address process variability, in particular concerning changes in the process correlation structure. In fact, only a comparatively low number of methodologies were proposed for explicitly addressing the problem of detecting changes in correlation.

The monitoring of the process dispersion and correlation is also a relevant issue since a process failure may not manifest itself so notoriously as a deviation from the nominal mean values, especially due to the action of control systems fighting to maintain key process variables close to their target values. In multivariate processes the monitoring of process dispersion is usually achieved by monitoring statistics based on the generalized variance [6–8] or through the implementation of supervision schemes based on successive likelihood ratio tests [4,9–11]. Yet, these procedures tend to consider only the marginal covariance and ignore the variables inner associations. Even the SPC methodologies based on state-space models (such as CVA and subspace system identification),

* Corresponding author. Tel.: +351 239 798 700; fax: +351 239 798 703.
 E-mail address: marco@eq.uc.pt (M.S. Reis).

are based on the marginal cross-covariance between the variables [12–14]. On the other hand, partial correlations have the potential to describe such local associations, even though in a non-causal way, being already applied to retrieve the structural information underlying collected data [15–17]. Moreover, Rato and Reis [18] showed recently that monitoring statistics based on partial correlations have indeed the potential to enhance the detection of structural changes, when compared to current state of the art methods. In this preliminary work [18], the authors established the base approach to monitor structural changes through the application of partial correlations and concluded that the proposed RMAX statistic (maximum norm of normalized partial correlations) was the one presenting better characteristics for detecting drifts in the process structure. Furthermore, it was also proved that structural changes are easier to detect for uncorrelated variables. These were obtained through a Cholesky decomposition, following a similar approach to Hawkins and Maboudou-Tchao [19]. Even though this approach can lead to a significant increase in detection performance in some circumstances, it does not consider the inner relationship between the variables. Therefore, the potential of effectively using the process correlated network to decorrelate the variables and for monitoring the process auto- and cross-correlation structure was not fully exploited yet. In the current work we explore this possibility and propose a new class of variables transformations based on the network reflecting the inner associations between variables which are obtained through the analysis of partial correlations, and that can finally lead to consistent superior detection performances. Additionally, a monitoring statistic specifically designed to detect changes in the variance will also be proposed here, in order to complement the monitoring of partial correlations which are insensitive to changes on the variables dispersion. Departures from the system's normal operation conditions (NOC) model can also be detected by this new monitoring statistic. Therefore, used in combination with the previously proposed RMAX statistic, it becomes possible to detect a full spectrum of faults specifically related with changes in the process structure.

Some related approaches have already been proposed mostly for fault diagnostics, based on the analysis of fault propagation pathways over an existent causal network [20,21]. However, these applications are dedicated to the diagnostic stage, assuming the existence of a previous detection phase and also the availability of a priori knowledge regarding variables connectivity. The same applies to model based approaches to fault detection, where a process dynamic model is always assumed to be available [22–24].

This article is organized as follows. In the next section, the current monitoring statistics based on the marginal covariance are briefly reviewed. Then, we describe the proposed monitoring statistics based on partial correlations for monitoring the process structure. The set of proposed sensitivity enhancing transformations studied are introduced in the next section. We then conduct a performance comparison study encompassing both the current monitoring statistics and the ones proposed in this article when applied to systems with different degrees of complexity, and present and discuss the results obtained. Finally, we summarize the main contributions proposed in this paper and present our conclusions.

2. Statistical monitoring of process structure based on marginal covariance

The procedures for monitoring the multivariate process dispersion are either based on the implementation of a likelihood ratio test or on the generalized variance and can be traced back to works

of Alt [6] and Alt and Smith [9]. In these articles, the authors propose to monitor the statistic,

$$W = -p(n-1) - (n-1) \ln \left(\frac{|\mathbf{S}|}{|\boldsymbol{\Sigma}_0|} \right) + (n-1) \text{tr}(\boldsymbol{\Sigma}_0^{-1} \mathbf{S}) \quad (1)$$

where, p stands for the number of variables, n is the number of observations, $\boldsymbol{\Sigma}_0$ is the in-control covariance matrix and \mathbf{S} is the sample covariance matrix. This monitoring statistic responds well to perturbations on low dimensional linear stationary systems and has the advantage of being invariant to any non-singular transformation of the data [19]. However, as it requires the inversion of the in-control covariance matrix, it may be unsuitable for cases where $\boldsymbol{\Sigma}_0$ is ill-conditioned, as happens in moderate/high dimensional processes.

Another procedure, based on the likelihood ratio test, was proposed by Levinson et al. [10] and explores the idea that under normal operation conditions the covariance matrix estimated with the complete data set should be equal to the one estimated by the mean square of successive differences. In order to obtain the monitoring statistics one has to calculate first the in-control sample covariance matrix, \mathbf{S}_0 , from a reference data set with n_0 observations. After that, the i th test sample covariance matrix, $\mathbf{S}_{1,i}$, determined from a subgroup with n_1 observations is combined with \mathbf{S}_0 in the pooled estimator of the covariance [10],

$$\mathbf{S}_{w,i} = \frac{(n_0-1)\mathbf{S}_0 + (n_1-1)\mathbf{S}_{1,i}}{n_0 + n_1 - 2} \quad (2)$$

The monitoring statistic is then computed as $G_i = m M_i$ where,

$$M_i = (n_0 + n_1 - 2) \ln |\mathbf{S}_{w,i}| - (n_0 - 1) \ln |\mathbf{S}_0| - (n_1 - 1) \ln |\mathbf{S}_{1,i}| \quad (3)$$

$$m = 1 - \left(\frac{1}{n_0 - 1} + \frac{1}{n_1 - 1} - \frac{1}{n_0 + n_1 - 2} \right) \left(\frac{2p^2 + 3p - 1}{6(p+1)} \right)$$

To specifically monitor the sample variance, a more efficient monitoring scheme proposed by Costa and Machado [25] can be applied. Their procedure is based on the sample covariance matrix, by taking the maximum value of the sample variance of normalized data, i.e.,

$$VMAX = \max\{s_1^2, s_2^2, \dots, s_p^2\} \quad (4)$$

where $s_i^2 = \mathbf{z}_i^T \mathbf{z}_i / n - 1$ and $\mathbf{z}_i = (\mathbf{x}_i - \boldsymbol{\mu}_i) / \sigma_i$. The normalization of the data is an important step on this procedure since it guarantees that all the sample variances have the same probability to exceed a certain UCL. Even though this procedure is simple and efficient, by design it is only capable to detect increases in variance.

3. Statistical monitoring of process structure based on partial correlations

Hitherto, the current monitoring statistics are typically based on measures of multivariate dispersion based on marginal information. It is well-known that such information leads to correlations between variables that are not directly associated, as long as they share some common inducing variation sources or are part of the same causality chain. Therefore, these methodologies are in principle unable to detect subtle and localized changes on the process structure. One way to improve the detection capability to faults specifically related with the local process structure relies on the use of partial correlations. Partial correlations have the ability to identify variables that are directly associated or connected, a feature that was already extensively explored to reconstruct interaction networks [15–17,26] as well as in classification problems [27–29]. In the SPC context, it is expected that monitoring partial correlations can bring the benefits of increasing sensitivity to more subtle changes in the process structure (due to the finer description of

the NOC structure) and to speed up the diagnosis stages (due to the localized nature of the information they provide), which potentially leads to a significant reduction in the total time to detect, diagnose and accommodate or fix, a given process upset. With these aims in mind, in the following subsections we present statistics for monitoring the fine process structure based on the information provided by partial correlations, along with their complementary statistics dedicated to monitor changes in variance. As will be demonstrated in the results section, the combined use of these complementary statistics can significantly increase the methods detection ability to a large class of faults.

3.1. ROMAX and R1MAX

The main idea behind the use of partial correlations is to bring out differences between directed and undirected variable's relationships by removing the common variables effects [27–29]. This is usually done by conditioning upon (i.e., controlling for, or holding constant) one or several other variables (z) before checking an association between the two designated variables (x and y). For such, one can adopt a linear regression approach where x and y are the responses and z (conditioning variables) the regressors, and then compute the correlation between the residuals obtained in both models for x and y . This correlation is known as the partial correlation ($r_{xy.z}$) of the original variables and is a measure of the level of direct association between the components of x and y that are uncorrelated with z . By conditioning over all the remaining variables, it is possible to infer the existence or not of a direct relationship between the pair under analysis, from which a finer and more detailed picture of the true variables interaction network can be obtained. Even though no causal directionalities are established at this point, the direct variables associations can be properly identified using this type of procedure. Instead of fitting multiple models for each pair of variables, the following recursive formulas (Eqs. (5)–(7)) can also be applied for computing the partial correlation coefficients of lower orders (the order of a partial correlation coefficient is given by the number of variables conditioned in its computation). Similar equations exist for higher orders.

0th order partial correlation:

$$r_{xy} = \frac{\text{cov}(x, y)}{\sqrt{\text{var}(x)\text{var}(y)}} \quad (5)$$

1st order partial correlation:

$$r_{xy.z} = \frac{r_{xy} - r_{xz}r_{yz}}{\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}} \quad (6)$$

2nd order partial correlation:

$$r_{xy.zq} = \frac{r_{xy.z} - r_{xq.z}r_{yq.z}}{\sqrt{(1 - r_{xq.z}^2)(1 - r_{yq.z}^2)}} \quad (7)$$

Following a reasoning analogous to network reconstruction methodologies, Rato and Reis [18] proposed to test partial correlations against their NOC distributions. The procedure consists of first pre-processing and normalizing the partial correlations to a standard normal distribution with zero mean and unit variance, and then, the maximum norm, $\|x\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_n|\}$, (i.e., the maximum of the absolute values) of these normalized partial correlations is considered as the monitoring statistic. This procedure is applied to both 0th order partial correlation (corresponding to the marginal correlation coefficients) and 1st order partial correlations. The monitoring statistics are defined as,

$$\text{ROMAX} = \|w(\mathbf{r}_0)\|_\infty = \max\{|w(\mathbf{r}_0)|\} \quad (8)$$

$$\text{R1MAX} = \|w(\mathbf{r}_1)\|_\infty = \max\{|w(\mathbf{r}_1)|\} \quad (9)$$

where $w(\cdot)$ is a normalization function based on the partial correlations distribution, \mathbf{r}_0 is the $(p(p-1)/2) \times 1$ column vector containing all distinct correlation coefficients (0th order partial correlations) and \mathbf{r}_1 is the $(p(p-1)(p-2)/2) \times 1$ column vector of 1st order partial correlation coefficients. The normalization function is easy to construct when the number of samples (n) is large, since the q th order partial correlation coefficients, tend to be normally distributed with zero mean and unit variance, when transformed according to [30],

$$w_1 = \frac{\sqrt{n-q-1}(r-\rho)}{1-\rho^2} \quad (10)$$

where r is the sample partial correlation and ρ is the population partial correlation.

Another useful normalization function is based on the Fisher's z transformation [30], which tends more rapidly to normality,

$$w_2 = \frac{\sqrt{n-q-1}}{2} \left[\ln\left(\frac{1+r}{1-r}\right) - \ln\left(\frac{1+\rho}{1-\rho}\right) \right] \quad (11)$$

In cases where it cannot be assumed that the partial correlation coefficient follows a normal distribution, an estimation of the density function can be applied, for instance based on kernel density estimation, which can then be used to normalize the partial correlations by matching the empirical distribution quantiles with those of the standard normal distribution.

3.2. VnMAX

As stated above in this article, monitoring solely based on the information provided by partial correlation coefficients is not enough to detect all faults, as some events may pass undetected such as those affecting process variance. This is particularly relevant when transformed variables are used, since a model mismatch can then be translated into changes on the transformed variables' variance. In this context, the VMAX statistic proposed by Costa and Machado [25] (see Section 2) is the one offering more potential for dealing with this problem. However, as it is defined, it is only able to detect a significant increase in the variance. Of course, in some situations this may be all that is necessary, but in the present context we are interested in detecting *any* departure from normal operation conditions. Therefore, we propose a modification of this statistic based on the inference properties of the variance in a normal distribution.

The test statistic for the hypothesis test that a sample variance obtained from n observations of a random normally distributed variable is equal to σ_0^2 , is given by Montgomery,

$$\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2} \quad (12)$$

where χ_0^2 follows a chi-squared distribution with $(n-1)$ degrees of freedom [31]. The one sided-test is already performed with the VMAX statistic, for detecting an increase in the variance (Section 2). In order to set the conditions for maintaining a methodology consistent with RMAX and use a single control limit, the probability of exceeding symmetrical UCL and LCL must be the same, which can be easily achieved by taken the absolute value of a symmetric random variable. In this particular case, the Wilson–Hilferty transformation can be applied to the chi-squared distribution to obtain a normally distributed random variable [32]. The use of this transformation also facilitates future constructions of a combined statistic to monitor both variance and correlation coefficients, since correlation coefficients are already approximately normally distributed.

The Wilson–Hilferty transformation states that a chi-squared distribution with ν degrees of freedom is well approximated by:

$$\chi_{\alpha,\nu}^2 \approx \nu \left[z_{\alpha} \left(\frac{2}{9\nu} \right)^{0.5} + \left(1 - \frac{2}{9\nu} \right) \right]^3 \quad (13)$$

where z_{α} is the $100 \times \alpha\%$ upper quantile from the standard normal distribution.

Given this relation, we propose to use the following transformation:

$$w_s = \frac{(s^2/\sigma_0^2)^{1/3} - (1 - (2/(9(n-1))))}{\sqrt{2/(9(n-1))}} \quad (14)$$

which is approximately distributed as $N(0, 1)$. Consequently, the modified (normalized) VnMAX statistic becomes:

$$VnMAX = \|w_s(\mathbf{v})\|_{\infty} = \max\{|w_s(\mathbf{v})|\} \quad (15)$$

where \mathbf{v} is a $(p \times 1)$ column vector containing the variables' variance.

As a consequence of the above transformation, this modified version is able to detect both increases and decreases in the process variance. It is also capable to detect changes in the process structure, when any of the sensitivity enhancing transformations to be described in Section 4, are applied. As an alternative to Eq. (14), a non-parametric density estimation methodology can also be employed as referred in Section 3.1 for the case of correlation coefficients.

4. Sensitivity enhancing transformations (SET)

The detection of changes on (marginal or partial) correlation coefficients is highly dependent of their nominal values under normal operation conditions. For instance, when the intrinsic relationship between two highly correlated variables suffers a small deviation, their correlation coefficient remains almost unchanged. On the contrary, whenever two initially unrelated variables become correlated to some extent, their correlation changes more abruptly than in other situations. A demonstration of this phenomenon, as well as an example of its impact on the correlation coefficients can be found in Ref. [18]. This feature suggests that, in order to detect small changes on the structure, it is advantageous to pre-process data in order to obtain uncorrelated variables. A similar principle was applied by Hawkins [33] with the regressed-adjusted variables. Yet, in his work, the variables transformation is intended just to eliminate the contribution of preceding variables in order to better isolate drifts in the mean value and not to enhance the detection ability to changes in correlation. The use of uncorrelated variables has also been applied to monitor the marginal covariance. Nevertheless, these procedures only adopt such transformations for simplification purposes (e.g. [19,34]). Moreover, they tend to use the inverse of the covariance matrix, which may be ill-conditioned, or triangularization methods (like the Cholesky decomposition [35]), which may not decorrelate data in the best way. Still, the latter approach showed to be quite useful, provided that the variables are properly ordered [18], and therefore it will be briefly introduced next in order to better evaluate the performance of the sensitivity enhancing transformations proposed in this paper.

The Cholesky decomposition consists of a series of successive linear regressions where the i th variable is regressed onto the remaining $(i-1)$ preceding variables, based on the factorization of the covariance matrix, Σ , into a lower triangular matrix \mathbf{L} , such that,

$$\Sigma = \mathbf{L}\mathbf{L}^T \quad (16)$$

This matrix, can then be used to obtain uncorrelated variables with unit variance through the following transformation [35]:

$$\mathbf{u} = \mathbf{L}^{-1}(\mathbf{x} - \boldsymbol{\mu}) \quad (17)$$

To account for the dynamic and non-linear features of the systems, Rato and Reis [18] suggested the inclusion of time-shifted variables or polynomial terms (depending of the situation) in the data matrix. Exemplifying for the case of including time-shifted variables, these additional variables should be placed at the beginning of the extended data matrix, as follows:

$$\tilde{\mathbf{X}} = [\mathbf{X}(l) \ \cdots \ \mathbf{X}(1) \ \mathbf{X}(0)] \quad (18)$$

where $\mathbf{X}(j)$ is an $n \times p$ matrix of variables shifted j times into the past (i.e., with j lags), so past variables become the regressors of present variables.

After this step, the regular Cholesky decomposition is performed, resulting in a new set of uncorrelated variables according to Eqs. (19) and (20).

$$\tilde{\Sigma} = \tilde{\mathbf{L}}\tilde{\mathbf{L}}^T \quad (19)$$

$$\tilde{\mathbf{u}} = \tilde{\mathbf{L}}^{-1}(\tilde{\mathbf{x}} - \tilde{\boldsymbol{\mu}}) \quad (20)$$

Note that only the regression variables related to the present state are of interest (i.e., the last p variables in $\tilde{\mathbf{u}}$), since they are the ones that correspond to the residuals of the linear regression of the variables in the current time, onto those from the past. Since this last transformation can be seen as an extended version of the typical Cholesky decomposition, it will be referred as T_{ChExt} in the rest of this paper.

As stated earlier, the Cholesky decomposition, is essentially a triangularization method that regresses the i th variable onto the other $(i-1)$ variables that preceded it. Therefore, the sequence by which variables are included in the model is defined by their particular order of appearance in the data matrix. In this case, it is unlikely that such an arbitrary ordering scheme will provide the best description of the system structure. In fact, the more meaningful variable ordering results from placing the more important variables (i.e., variables that affect most of the others or that are placed on root nodes of the true variable interaction network) at the beginning, by considering a priori knowledge of the process. As this information may not be always available, an alternative transformation is proposed. The goal of the new transformation is to break the relevant variables relationships upon application of linear regression only on the variables that are indeed related, as exemplified on Fig. 1 ("breaking" means here the removal of an association by application of a proper model). For such, the relevant edges between variables must be first identified through a network reconstruction technique, as described in the next paragraph (see also Fig. 1(a) and (b)).

In order to identify the edges linking directly associated variables, we suggest the use of partial correlation up to the 2nd order, according to a procedure similar to that applied elsewhere [16,28,29]. The main difference relies on the inclusion of time-shifted variables or polynomial terms in the regression models, for addressing dynamic and non-linear process features, respectively. By conditioning on these additional variables, this algorithm becomes capable to detect both dynamic and non-linear relationships, resulting in more accurate networks. After that, the causal directions of these undirected edges are determined in a second stage, by applying another algorithm based on the variables' cross-correlation. Alternatively, one can also use Granger causality [21] or transfer entropy [20] to determine the variables causal direction. The pseudo-code for these algorithms, are available in the journal website as supplementary material.

The second step of the transformation involves the regression of each variable onto its parents (see Fig. 1(c)), resulting in a final

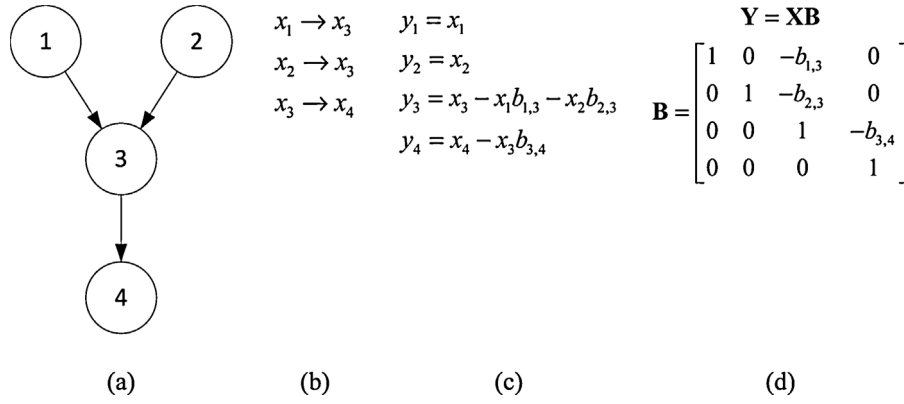


Fig. 1. Implementation of the proposed transformation procedure: (a) original network, (b) identification of the relevant edges, (c) implementation of successive linear regressions involving the directly connected variables and (d) final model.

Table 1

Nomenclature associated with the sensitivity enhancing transformations covered in this study.

Transformation	Description
T_X	Original variables
T_{ChExt}	Variables transformed using the Cholesky decomposition after incorporation of time-shifted variables and/or polynomial terms, Eq. (20)
T_{Net}	Variables transformed using a set of linear regression models constructed only for related variables (see Fig. 1)
T_{NetCh}	Variables transformed using a set of linear regression models constructed only for related variables with an additional Cholesky decomposition of the resulting residuals

regression model where only the directly connected variables are considered to obtain the new set of residual variables. On dynamic systems, time-shifted variables should also be included on the regression model and for non-linear systems, polynomial terms must also be added. As a possible additional step, a Cholesky decomposition can be applied to the residuals, in order to ensure that uncorrelated variables are indeed obtained and also to accommodate any missed relationship. On the rest of the article, these novel NETWORK-oriented transformations will be referred as T_{Net} (without a Cholesky decomposition of the residual) and T_{NetCh} (with a Cholesky decomposition of the residual).

These new proposed transformations are able to integrate information about the local structure of correlations, leading to improvements in the monitoring statistics performance. Furthermore, as the variables connectivity is estimated through network reconstruction methods (using the algorithms proposed here), the nature of the transformations remain fully data driven.

The complete set of sensitivity enhancing transformations covered in this study, and the associated nomenclature, are summarized in Table 1. The proposed monitoring statistics for each transformation are presented in Table 2.

5. Results

The effects of the proposed sensitivity enhancing transformations in the performance of the developed monitoring statistics is here compared against their current counterparts already proposed in the literature, namely the ones described in Section 2. Other monitoring statistics were also considered, but excluded from this presentation due to their lower relative performance. This analysis was performed for three simulated systems with different

Table 2

Definition of the monitoring statistics proposed in this work.

Sensitivity enhancing transformation	Statistic		
	$VnMAX$	$RMAX$	
	\mathbf{v}	\mathbf{r}_0	\mathbf{r}_1
T_X	$VnMAX_X$	$ROMAX_X$	$R1MAX_X$
T_{ChExt}	$VnMAX_{ChExt}$	$ROMAX_{ChExt}$	$R1MAX_{ChExt}$
T_{Net}	$VnMAX_{Net}$	$ROMAX_{Net}$	$R1MAX_{Net}$
T_{NetCh}	$VnMAX_{NetCh}$	$ROMAX_{NetCh}$	$R1MAX_{NetCh}$

degrees of complexity in order to demonstrate the versatility of the proposed methodology, its robustness to misspecifications of the causal structure and the consistently higher detection capability regarding other alternative frameworks.

For each case study, faults of different magnitudes were generated and the associated curve of detection rate versus fault's magnitude was determined. For the sake of making it possible to establish a sound comparison between the different methods, each curve is summarized by a performance index (N) based on the area under the curve. Note that, as the fault's magnitude increases, the detection rate also increases, until a maximum value of 1 (i.e. all points are out-of-control). Therefore, a monitoring statistic with a greater area under the detection rate curve tends to 1 more rapidly and will present higher detections rates. This index was computed for each fault and normalized so that its values fall in the range [0, 1], where 1 represents the best performance observed in each fault, by a given method. In this way the relative performance of different methods in each fault can be established and easily analyzed. Additionally, we have also assessed the statistical significance of the detection rates of the studied monitoring statistics through the application of a permutation test [36], from where the associated p -values were determined. These results are available in the journal website as supplementary material.

5.1. Case study 1: stationary non-linear system

To test the performance of the proposed monitoring methods (SET and statistics), the causal network proposed by Tamada et al. [37] was adopted. This network is composed by 16 nodes (or variables) causally related according to the representation provided in Fig. 2. The variables relationships were approximated by polynomials according to Eq. (21), where ε_i is a white noise sequence with a signal-to-noise ratio of 10 dB. By application of the identification and causal directionality inference algorithms referred in Section 4 to a data set with NOC data, the original causal network was obtained (full reconstruction). The variable transformation T_{Net} was then carried out taken this inferred structure into account, and

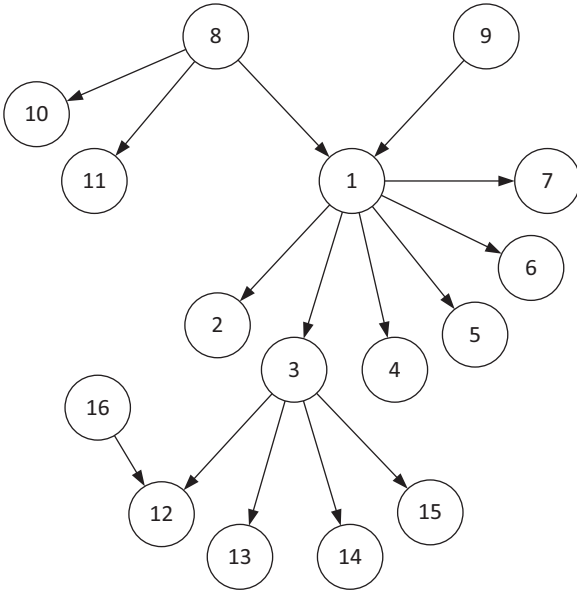


Fig. 2. Graphical representation for the causal structure for the artificial network used in case 1.

Table 3

Definition of the faults and respective variables involved for the network system with a non-linear model structure. δ is a multiplicative factor that changes the model parameters (under NOC, $\delta = 1$).

Fault	Variables relation changed
A	$g_8 \rightarrow g_1 (g_1 = 1.20\delta g_8 + 0.80g_9 + \varepsilon_1)$
B	$g_1 \rightarrow g_3 (g_3 = (\delta g_1 - 4)(g_1 + 4) + \varepsilon_3)$
C	$g_8 \rightarrow g_{10} (g_{10} = 0.02g_8^2 + 0.44\delta g_8 + 0.82 + \varepsilon_{10})$
D	$g_3 \rightarrow g_{14} (g_{14} = 0.020\delta g_3^2 + 0.44g_3 - 0.82 + \varepsilon_{14})$
E	$g_8 \rightarrow g_{11} (g_{11} = -0.053\delta g_8^3 - 0.00068g_8^2 + 0.52g_8 + 0.50 + \varepsilon_{11})$
F	$g_3 \rightarrow g_{15} (g_{15} = 1.40\delta g_3 + \varepsilon_{15})$

a 3rd order polynomial model was finally fitted for each variable using the identified variables' parents as regressors (for instance, in the case of variable 1 the regressors identified were variables 8 and 9). Transformation T_{ChExt} was also obtained from the same reference data set by adding polynomial terms to the data matrix. These transformed variables, along with the original ones, were then used as inputs of the monitoring statistics of the process structure.

The system was perturbed with the faults presented in Table 3, where the magnitude of change, δ , was set to cause changes on the model's parameter on the range of $\pm 10\%$. In each perturbation, 1000 sample covariance matrices were determined with 3000 observations each. Based on the sample covariance matrices, we determined the monitoring statistics and their corresponding fault detection rates. The same procedure was repeated 10 times in order to estimate the confidence levels of the detection rates. The control limits for all the monitoring statistics were preliminarily adjusted so that all monitoring statistics present the same false detection rate of 1%. Subsequently, for each fault the performance index N was obtained, resulting in the global distribution, as shown in Fig. 3. Similarly, the results of the permutations tests are summarized in Table S5 of the supplementary material.

Analyzing the results obtained, it is possible to observe that none of the current monitoring statistics presented detections rates higher than 5%. This result is a clear evidence that these monitoring statistics are not able to deal with the systems' features and stress the need for more efficient monitoring schemes. The proposed statistics, RMAX and VnMAX, were the only ones capable to effectively detect these faults. Furthermore, it is also evident that not all of the sensitivity enhancing transformations

are appropriate. For instance, T_{ChExt} does not model correctly the systems non-linearity, leading to partial correlations that do not follow their assumed NOC probability distributions. Therefore, the RMAX statistic cannot be directly applied with this transformation, since RMAX requires that all partial correlation coefficients have the same probability of exceeding the control limits. The same situation arises with the original untransformed variables. Nevertheless, $RMAX_{ChExt}$ and $R1MAX_{ChExt}$ still present some capabilities to detect faults and when the partial correlations are properly pre-processed with resource to an adequate estimation of their distribution, the RMAX performance is significantly increased. Yet, the use of the proposed transformation T_{Net} is preferable because the actual variables' relationships are explicitly considered in its construction. Transformation T_{Net} also leads to uncorrelated variables and partial correlations following normal distributions. Consequently, $RMAX_{Net}$ and $R1MAX_{Net}$ can detect efficiently most of the simulated faults, with the exception of fault D which is only detected by $VnMAX_{Net}$. Hence, transformation T_{Net} is adequate for modelling and monitoring this non-linear system, since either $RMAX_{Net}$ or $VnMAX_{Net}$ are able to rapidly detect any structural change, while current monitoring statistics in general fail in this task.

$$g_8 = \varepsilon_8, g_9 = \varepsilon_9, g_{16} = \varepsilon_{16}$$

$$g_{10} = 0.020g_8^2 + 0.44g_8 + 0.82 + \varepsilon_{10}$$

$$g_{11} = -0.053g_8^3 - 0.00068g_8^2 + 0.52g_8 + 0.50 + \varepsilon_{11}$$

$$g_1 = 1.20g_8 + 0.80g_9 + \varepsilon_1$$

$$g_2 = 0.60g_1 + \varepsilon_2$$

$$g_3 = (\delta g_1 - 4)(g_1 + 4) + \varepsilon_3$$

$$g_4 = 0.020g_1^2 + 0.44g_1 + 0.82 + \varepsilon_4$$

$$g_5 = -0.057g_1^3 - 0.077g_1^2 + 0.52g_1 + 0.22 + \varepsilon_5$$

$$g_6 = 0.60g_1 + \varepsilon_6$$

$$g_7 = 0.70g_1 + \varepsilon_7$$

$$g_{12} = 0.80g_{16} + 0.60g_3 + \varepsilon_{12}$$

$$g_{13} = 1.30g_3 + \varepsilon_{13}$$

$$g_{14} = 0.020g_3^2 + 0.44g_3 - 0.82 + \varepsilon_{14}$$

$$g_{15} = 1.40g_3 + \varepsilon_{15}$$

5.2. Case study 2: gene network model

The test system considered in this case study was proposed by Fuente et al. [16]. It regards the dynamic modelling of transcription levels (T_i) in a gene network, and consists of the following set of differential equations:

$$\begin{aligned} \frac{dT_1}{dt} &= \frac{V_1}{(1 + K_{T_1}/T_4)} - k_1 T_1 + \theta_1 T_1 \\ \frac{dT_2}{dt} &= \frac{V_2}{(1 + T_1/K_{T_1})} - k_2 T_2 + \theta_2 T_2 \\ \frac{dT_3}{dt} &= \frac{V_3}{(1 + K_{T_2}/T_2)} - k_3 T_3 + \theta_3 T_3 \\ \frac{dT_4}{dt} &= \frac{V_4}{(1 + T_2/K_{T_2})} - k_4 T_4 + \theta_4 T_4 \\ \frac{dT_5}{dt} &= \frac{V_5}{(1 + T_3/K_{T_3})(1 + K_{T_4}/T_4)} - k_5 T_5 + \theta_5 T_5 \end{aligned} \quad (22)$$

where V_i are maximal transcription rates, k_i degradation rate constants, K_i inhibition or activation constants and θ_i error terms designed to simulate biological variability. Parameters V_i , k_i and

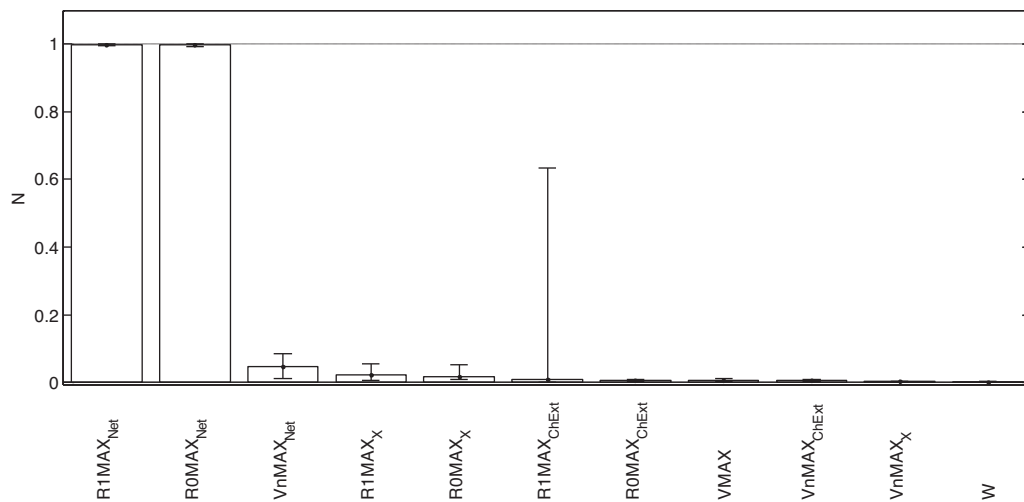


Fig. 3. Comparison of the statistics performance for case study 1 (network system with a non-linear model structure): performance index N based on the area under the fault detection curve obtained for all perturbations. Bars correspond to the associated median bounded by the 25th and 75th percentiles.

K_i were set to unity and the error terms (θ_i) were sampled from a normal distribution with zero mean and standard deviation of 0.01. The underlying network is represented in Fig. 4. The recovery of the undirected network from data was already addressed by Fuente et al. [16] through the use of partial correlations. As the directions of the edges can also be obtained by simple analysis of the cross-correlation between the variables, the complete network was considered in the construction of transformations T_{Net} and T_{NetCh} .

The system was subject to perturbations on V_1 , V_5 , k_1 and k_5 in a magnitude range of $\pm 20\%$. The control limits for the monitoring statistics were adjusted to a false detection rate of 1% under normal operation conditions. Afterwards, 500 sample covariance matrices were computed from 3000 observations from which the detection rates for each fault were determined. This procedure was repeated

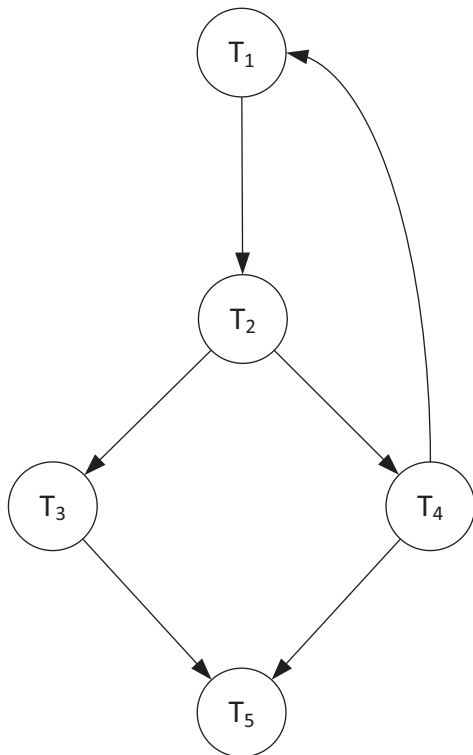


Fig. 4. Directed network representation of the gene network model.

5 times in order to obtain approximate confidence intervals for the fault detection rates. From these results, the performance index N presented in Fig. 5 was obtained. The results for the permutation tests are available in Table S6 of the supplementary material. Analyzing these results, it is possible to verify that, for this particular system, the current monitoring statistics presented higher detection rates than RMAX based statistics, regardless of the transformation used. This can be explained by analysis of the original untransformed data, where only small changes in the correlation matrix are observed. In fact, the most significant changes occurred in the process mean, since the faults drive the process to a new steady state, and in a lesser extent in the process' variance. Therefore, as none of the partial correlation based statistics account for changes in variance, all of them show poor detection performances under these simulated scenarios.

For detecting a change on the variance, the VMAX statistic was extended to detect both increases and decreases in process variance. Its direct extension without transformation, VnMAX_X, showed to be slightly worse than VMAX in detecting increases in variance, due to the increase in the control limit. However, when the variance decreases, VnMAX_X was capable to signal such change while VMAX is unable to detect it at all. Consequently VnMAX_X is globally better than VMAX when the detection of both increases and decreases are of interest, as can be seen in Fig. 5. The performance of VnMAX can be further improved by use of an appropriated transformation, namely transformations T_{ChExt} , T_{Net} and T_{NetCh} , which incorporate dynamic dependencies. For this particular case, the relative performance of VnMAX_{ChExt}, VnMAX_{Net} and VnMAX_{NetCh} is very similar (see Fig. 5 and Table S6). This happens because the order of the variables used to implement the transformation T_{ChExt} is consistent with their causal relationships (see Fig. 4) and therefore the model obtained is adequate. However, in the cases where any other variable ordering is used, the performance of VnMAX_{ChExt} changes, while that for VnMAX_{Net} and VnMAX_{NetCh} remains the same since they are not dependent on the particular order of appearance of the variables.

We would like to point out that perturbations detected by VnMAX are not only related to changes in the variance, but also to changes in the process' structure when transformed variables are used. In this case, if a change in the process' structure occurs, the transformation no longer describes the variables' relationships correctly. Therefore, non-accounted contributions will lead to deviations on the transformed variables' variance, which will ultimately be captured by VnMAX. This behaviour shows the importance of

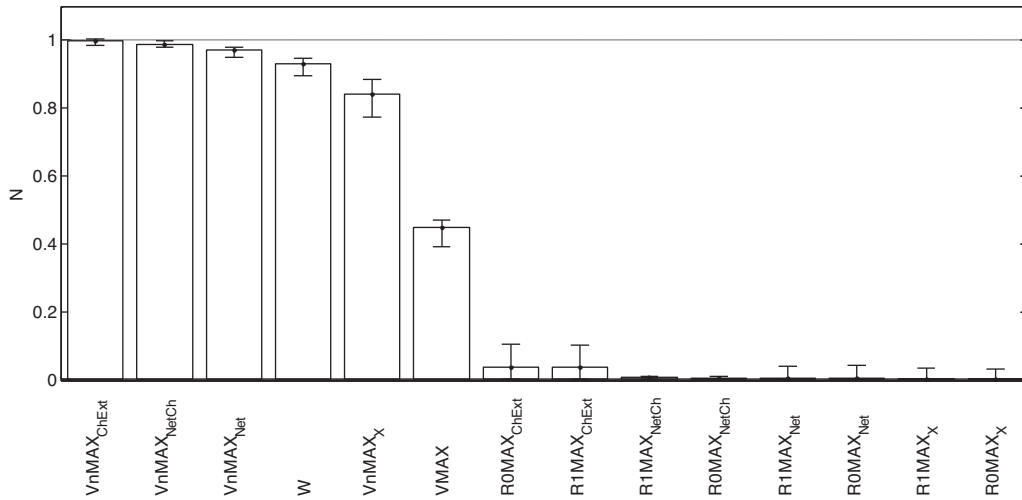


Fig. 5. Comparison of the statistics performance for case study 2 (gene network model): performance index N based on the area under the fault detection curve obtained for all perturbations. Bars correspond to the associated median bounded by the 25th and 75th percentiles.

Table 4
Jöbjes' model parameters.

Parameters	Values	Unit
K_E	0.00383	$\text{m}^6 \text{kg}^{-2} \text{h}^{-1}$
c_1	59.2085	kg m^{-3}
c_2	70.5565	kg m^{-3}
K_S	0.500	kg m^{-3}
m_S	2.160	$\text{kg kg}^{-1} \text{h}^{-1}$
m_P	1.10	$\text{kg kg}^{-1} \text{h}^{-1}$
Y_{SX}	0.02445	kg kg^{-1}
Y_{PX}	0.05263	kg kg^{-1}
μ_{\max}	1.0	h^{-1}

monitoring the variance in order to also detect structural changes, when using SET.

5.3. Case study 3: biologic production of ethanol

This non-linear dynamical system is based in the ethanol production from glucose fermentation by *Zymomonas mobilis* bacteria. One of the models proposed to describe the dynamics of this process is the Jöbjes' model [38], given by the following equations:

$$\begin{aligned}
 \frac{dC_S}{dt} &= - \left(\frac{\mu_{\max} C_S C_E}{Y_{SX}(K_S + C_S)} \right) - m_S C_X + D(C_{S0} - C_S) \\
 \frac{dC_X}{dt} &= \left(\frac{\mu_{\max} C_S C_E}{K_S + C_S} \right) + D(C_{X0} - C_X) \\
 \frac{dC_E}{dt} &= K_E(C_P - c_1)(C_P - c_2) \left(\frac{C_S C_E}{K_S + C_S} \right) + D(C_{E0} - C_E) \\
 \frac{dC_P}{dt} &= - \left(\frac{\mu_{\max} C_S C_E}{Y_{PX}(K_S + C_S)} \right) - m_P C_X + D(C_{P0} - C_P)
 \end{aligned} \quad (23)$$

where C_S is the substrate (glucose) concentration, C_X is the biomass (*Zymomonas mobilis*), C_P is the product (ethanol) concentration, and C_E is an auxiliary variable used to account for the lagged effect of ethanol concentration in the kinetic model. The variables C_{S0} , C_{X0} , C_{E0} and C_{P0} complete the mass balance representing the input concentrations in the reactor, where normally only C_{S0} (substrate inlet) is non-zero. The dilution rate (D) is the inverse of the resident time. The reactor volume was kept constant and all the reminding parameters proposed by Jöbjes et al. [38] are listed in Table 4. For our case study, only C_S , C_X and C_P were considered to be measured variables.

By analysis of the differential equations, the system model represented by Eq. (23), can be translated into the causal network presented in Fig. 6(a), by considering that if a variable is present in the right side of the differential equation, then it has an effect on the main variable (that appears on the left side). This representation clearly shows a complex relationship between variables, as they are all directly or indirectly linked by closed-loops. The reconstruction of the correct dependency between the variables connective structure has an impact in the monitoring statistics, since the derivation of the sensibility enhancing transformations depends on this map.

All the variable transformations described in Section 4 were determined based on a data set collected under normal operation conditions. In what concerns to transformation T_{ChExt} , no particular order was considered for the variables. As for transformations T_{Net} and T_{NetCh} , the algorithm developed for identifying the variables edges and dependencies led to the graph shown in Fig. 6(b). The differences between the reconstructed network and the real one (presented in Fig. 6(a)), lie essentially in the fact that only one direction is considered together with the absence of the unmeasured variable, C_E . However, even with these structural mismatches, the transformation is able to break the major cross- and auto-correlations present in the original data, and thus it provides a reasonable description of the system for monitoring purposes.

In order to compare the monitoring statistics performance, the system was subject to changes on c_1 , K_S and Y_{SX} . The range of the perturbations was of $\pm 25\%$ for c_1 and Y_{SX} and $\pm 50\%$ for K_S . The control limits of the studied monitoring statistics were set to a false

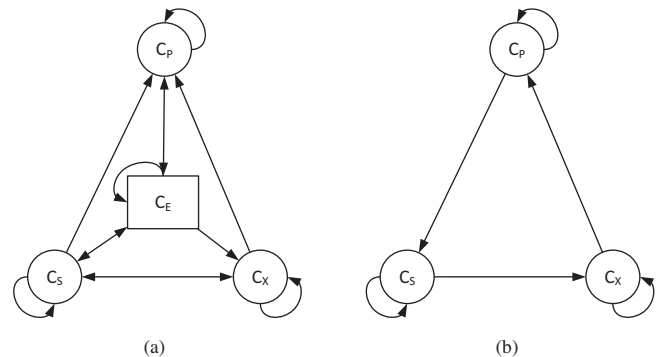


Fig. 6. Jöbjes' model network: (a) the original causal network, (b) the estimated causal network. Circles represent measured variables and rectangles unmeasured variables.

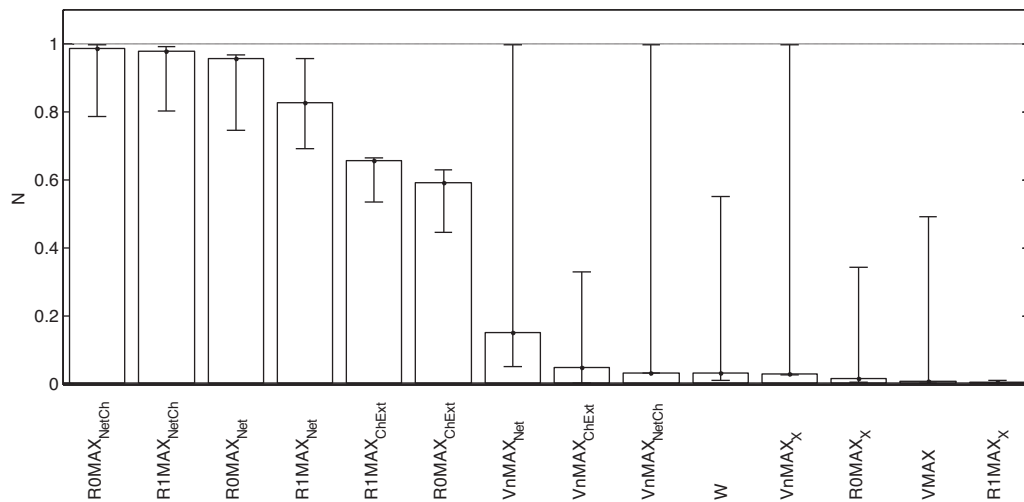


Fig. 7. Comparison of the statistics performance for case study 3 (Jöbess' model): performance index N based on the area under the fault detection curve obtained for all perturbations. Bars correspond to the associated median bounded by the 25th and 75th percentiles.

detection rate of 1% and the detection rates were determined from 500 sample covariance matrices based on 3000 observations each. These simulations were repeated 5 times in order to assess the results consistency. As in the previous cases study, the performance index N (Fig. 7) and permutations tests (Table S7 in supplementary material) were computed and saved for analysis.

Looking to the current monitoring statistics, the W and G statistics were the ones that presented the best performance. However none of the current monitoring statistics successfully detects faults on K_S and Y_{SX} , leading to fault detections rates below 5%. On the other hand, most of the proposed monitoring statistics are able to signal all the simulated faults. Only the statistics based on the original variables (T_X) presented poor detection capabilities. This result is expected as the T_X variables are highly correlated, and therefore changes are difficult to detect. The effect of cross- and auto-correlation is eliminated through the use of the transformation T_{ChExt} , which includes time-shifted variables on the Cholesky decomposition. This is the simplest among the three studied sensitivity enhancing transformations for dynamic systems and consequently both T_{Net} and T_{NetCh} transformations present even better performances, since they make use of more information regarding the variables' relationships.

The main difference between transformations T_{Net} and T_{NetCh} is the additional Cholesky decomposition performed over the final residuals in the case of T_{NetCh} . This extra step ensures that the variables obtained are indeed uncorrelated and it accommodates any relationship that is not being considered by T_{Net} . This characteristic gives some robustness to the proposed transformations, since it allows some flexibility in the estimation of the causal network. As stated before, the retrieved causal network for this case study is substantially different from the real one. Yet, with transformation T_{NetCh} , the effects of the missed interactions are mitigated and an increase in the monitoring statistics performance is observed (see Fig. 7). This shows that the proposed monitoring statistics based on this type of SET presents an interesting degree of robustness, a useful feature in real world practical applications.

6. Discussion

The statistical process monitoring of process multivariate dispersion is mostly done by analysis of either the generalized variance, likelihood test ratio or other dispersion measures that only consider the marginal distributions of process data. These descriptions have a rather small resolution of the process inner

structure and therefore such procedures are unable to effectively detect localized changes. This feature becomes more relevant when the system cannot be assumed to be approximately linear around its operation state, as in the case study considered in Section 5.1.

In order to incorporate the systems' non-linearity and dynamics in a more rigorous and consistent way, we proposed new sensitivity enhancing transformations that use partial correlations to retrieve the underlying process structure. The application of these transformations lead to a new set of uncorrelated variables that account for the process structure and, due to their uncorrelated nature, makes fault detection easier. Moreover, as the system is compared against a fix model, any drift from the reference model can also be perceived as a structural change. These types of faults are essentially manifested in changes in variance and can be detected by an extended version of the current monitoring statistics VMAX. The proposed extension, VnMAX, is capable to detect both increases and decreases in variance, complementing the RMAX statistics in such a way that any structural change is detected by at least one of them. Therefore, the combined use of these statistics will enable to monitor the full spectrum of structural changes conveyed by both VnMAX and RMAX. Yet, on this study, they were treated separately to clarify and understand their complementarity.

Another important feature of the sensitivity enhancing transformations is their ability to increase the sensitivity to changes in correlation. This increase in sensitivity occurs because around the state of zero correlation, any small change in the variables relationship is translated into a high deviation on the correlation coefficient values. Furthermore, as time-shifted variables and/or polynomial terms are added to the model, the transformations also break the systems dynamics and non-linearity, facilitating the detection of faults related to these terms. Note that the conventional approach only analyses the linear relationships between the variables and therefore cannot detect more complex relationships. Additionally, the proposed network-oriented transformations are far superior than the method based on the Cholesky decomposition proposed earlier, where the monitoring statistics performance is highly dependent on the variables order. However, the new transformations require the estimation of the causal network, which can be a quite complex task. Still, our study shows that when the regression step is complemented by an additional Cholesky decomposition over the regression residuals, the final model has some robustness to a miss specification of the true network, namely the identification of wrong edges and the incorrect attribution of a causal direction.

For multivariate processes, the need of fitting a regression model for each variable can be seen as a disadvantage. However, it is worth noticing that other similar approaches have already been proposed, such as Hawking's regressed-adjusted variables [33] or Ottestad's regression components [39]. Other examples can also be found in Refs. [40,41]. However, none of these methods are driven by the inner process structure, with the aim of maximizing the detection of correlation drifts. In this study, we applied Ordinary Least Squares regression (OLS), but Principal Component Regression (PCR) or Partial Least Squares Regression (PLSR) could have been applied as well to solve potential collinearity problems that may arise in the regression stage. Still, the resource to such regression methodologies was not necessary, since the proposed methodology pre-selects the set of variables to be used as regressors based on the inferred causal network of the process and these variables did not present collinearity issues.

The procedures mentioned above do not consider the inner process structure and even Principal Component Analysis (PCA), which is extensively used to explain the process variability, ends up with a linear combination of variables that may not be directly related. Moreover, the typical multivariate control schemes based on PCA, have some limitations in the detection of changes in the correlation and are easily outperformed by monitoring statistics specifically devoted for such task, as for instance the W statistic [18].

After application of an appropriate sensitivity enhancing transformation, the monitoring of the marginal correlation by R0MAX or of the 1st order partial correlations by R1MAX, seems to be statistically similar. This situation arises from the close relationship between these two quantities and also because the transformed variables are inherently uncorrelated, which means that there is no extra effect in controlling any pair of variables by a third one. Still, in both cases, the inner process structure is explored, since the transformed variables result from the application of a model based on partial correlation, which are well-known by their capability to uncover the variables relationships. Furthermore, the whole concept of partial correlations is based on regression residuals obtained by controlling two variables by a set of others. Consequently, the marginal correlation of the transformed variables can, to some extent, be considered as partial correlations of the original variables, since the transformed variables are in fact the regression residuals of the original variables controlled by their parent variables.

7. Conclusions

In this work we have addressed the implementation of sensitivity enhancing transformations in statistical process monitoring of the process structure by means of partial and marginal correlations. The use of partial correlations allows for a better description of the systems dynamics and non-linearity and consequently, a higher detection performance. Most of this improvement is due to the introduction of process structure information into the sensitivity enhancing transformation.

In order to expand the detection capability of faults related with the process structure, we have also proposed the VnMAX statistic. This monitoring statistic is a generalization of the current VMAX statistic, and it is able to detect both changes in variance and in the process structure, since model deviations lead to changes in the transformed variables variance. This result is another indication that the sensitivity enhancing transformations increase the sensitivity to detect changes in correlation, through RMAX, and simultaneously assess the model validity, through VnMAX. These two monitoring statistics can also be easily combined into a single monitoring scheme, able to detect both types of deviations and therefore consolidate the detection of structural changes through an unified monitoring scheme.

The proposed monitoring statistics also proved to be statistically superior when compared to the current monitoring statistics tested, which makes this approach a feasible solution worthwhile considering in practical applications. Future work will encompass the development of similar monitoring schemes for single observations, instead of the use of rational groups of observations.

Acknowledgments

Tiago J. Rato acknowledges the Portuguese Foundation for Science and Technology for his PhD grant (grant SFRH/BD/65794/2009). Marco S. Reis also acknowledges financial support through project PTDC/EQU-ESI/108374/2008 co-financed by the Portuguese FCT and European Union's FEDER through "Eixo I do Programa Operacional Factores de Competitividade (POFC)" of QREN (with ref. FCOMP-01-0124-FEDER-010397).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jprocont.2014.04.006>.

References

- [1] V.B. Ghute, D.T. Shirke, A multivariate synthetic control chart for process dispersion, *Qual. Technol. Quant. Manage.* 5 (3) (2008) 271–288.
- [2] A.B. Yeh, D.K.J. Lin, R.N. McGrath, Multivariate control charts for monitoring covariance matrix: a review, *Qual. Technol. Quant. Manage.* 3 (4) (2006) 415–436.
- [3] B. Abbasi, et al., A transformation-based multivariate chart to monitor process dispersion, *Int. J. Adv. Manuf. Technol.* 44 (7) (2009) 748–756.
- [4] C.-L. Yen, J.-J.H. Shiau, A.B. Yeh, Effective control charts for monitoring multivariate process dispersion, *Qual. Reliab. Eng. Int.* 28 (4) (2012) 409–426.
- [5] S. Bersimis, S. Psarakis, J. Panaretos, Multivariate statistical process control charts: an overview, *Qual. Reliab. Eng. Int.* 23 (5) (2007) 517–543.
- [6] F.B. Alt, Multivariate quality control, in: S. Kotz (Ed.), *Encyclopedia of Statistical Sciences*, John Wiley & Sons, Inc., New York, 2005, pp. 5312–5323.
- [7] M.A. Djuhari, Improved monitoring of multivariate process variability, *J. Qual. Technol.* 37 (1) (2005) 32–39.
- [8] F. Aparisi, J. Jabaloyes, A. Carrión, Generalized variance chart design with adaptive sample sizes. The bivariate case, *Commun. Stat. Simul. Comput.* 30 (4) (2001) 931–948.
- [9] F.B. Alt, N.D. Smith, Multivariate process control, in: P.R. Krishnaiah, C.R. Rao (Eds.), *Handbook of Statistics*, Elsevier, 1988, pp. 333–351.
- [10] W.A. Levinson, D.S. Holmes, A.E. Mergen, Variation charts for multivariate processes, *Qual. Eng.* 14 (4) (2002) 539–545.
- [11] C.-L. Yen, J.-J.H. Shiau, A multivariate control chart for detecting increases in process dispersion, *Stat. Sinica* 20 (2010) 1683–1707.
- [12] E.L. Russell, L.H. Chiang, R.D. Braatz, Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis, *Chemometr. Intell. Lab. Syst.* 51 (1) (2000) 81–93.
- [13] A. Negiz, A. Çinar, Statistical monitoring of multivariable dynamic processes with state-space models, *AIChE J.* 43 (8) (1997) 2002–2020.
- [14] R.J. Treasure, U. Kruger, J.E. Cooper, Dynamic multivariate statistical process control using subspace identification, *J. Process Control* 14 (3) (2004) 279–292.
- [15] J.-P. Pellet, A. Elisseeff, A partial correlation-based algorithm for causal structure discovery with continuous variables, in: *Proceedings of the 7th International Conference on Intelligent Data Analysis*, Springer-Verlag, Ljubljana, Slovenia, 2007, pp. 229–239.
- [16] A.D.L. Fuente, et al., Discovery of meaningful associations in genomic data using partial correlation coefficients, *Bioinformatics* 20 (18) (2004) 3565–3574.
- [17] M. Kalisch, P. Bühlmann, Robustification of the PC-algorithm for directed acyclic graphs, *J. Comput. Graph. Stat.* 17 (4) (2008) 773–789.
- [18] T.J. Rato, M.S. Reis, Non-causal data-driven monitoring of the process correlation structure: a comparison study with new methods (in submission).
- [19] D.M. Hawkins, E.M. Maboudou-Tchao, Multivariate exponentially weighted moving covariance matrix, *Technometrics* 50 (2) (2008) 155–166.
- [20] M. Bauer, et al., Finding the direction of disturbance propagation in a chemical process using transfer entropy, *IEEE Trans. Control Syst. Technol.* 15 (1) (2007) 12–21.
- [21] T. Yuan, S.J. Qin, Root cause diagnosis of plant-wide oscillations using granger causality, in: *8th IFAC International Symposium on Advanced Control of Chemical Processes*, Singapore, 2012.
- [22] P.M. Frank, Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy – a survey and some new results, *Automatica* 26 (3) (1990) 459–474.

- [23] R. Isermann, Model-based fault-detection and diagnosis – status and applications, *Ann. Rev. Control* 29 (1) (2005) 71–85.
- [24] R. Isermann, Process fault detection based on modeling and estimation methods – a survey, *Automatica* 20 (4) (1984) 387–404.
- [25] A. Costa, M. Machado, A new chart based on sample variances for monitoring the covariance matrix of multivariate processes, *Int. J. Adv. Manufact. Technol.* 41 (7) (2009) 770–779.
- [26] A. Reverter, E.K.F. Chan, Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks, *Bioinformatics* 24 (21) (2008) 2491–2497.
- [27] A.S. Melissa, K.R. Raghuraj, S. Lakshminarayanan, Partial correlation metric based classifier for food product characterization, *J. Food Eng.* 90 (2) (2009) 146–152.
- [28] K.R. Rao, S. Lakshminarayanan, Partial correlation based variable selection approach for multivariate data classification methods, *Chemometr. Intell. Lab. Syst.* 86 (1) (2007) 68–81.
- [29] R. Rao, S. Lakshminarayanan, Variable interaction network based variable selection for multivariate calibration, *Anal. Chim. Acta* 599 (1) (2007) 24–35.
- [30] T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed., Wiley, New Jersey, 2003.
- [31] D.C. Montgomery, *Introduction to Statistical Quality Control*, 6th ed., John Wiley & Sons, New York, 2008.
- [32] E.B. Wilson, M.M. Hilferty, The distribution of chi-square, in: *Proceedings of the National Academy of Sciences of the United States of America*, 1931, pp. 684–688.
- [33] D.M. Hawkins, Regression adjustment for variables in multivariate quality control, *J. Qual. Technol.* 25 (3) (1993) 170–182.
- [34] L. Huwang, A.B. Yeh, C.-W. Wu, Monitoring multivariate process variability for individual observations, *J. Qual. Technol.* 39 (3) (2007) 258–278.
- [35] W.H. Press, et al., *Numerical Recipes: The Art of Scientific Computing*, 3rd ed., Cambridge University Press, Cambridge, 2007.
- [36] F. Pesarin, L. Salmaso, *Permutation Tests for Complex Data: Theory, Applications and Software*, John Wiley & Sons Ltd., New York, 2010.
- [37] Y. Tamada, et al., Estimating gene networks from gene expression data by combining Bayesian network model with promoter element detection, *Bioinformatics* 19 (1) (2003) ii227–ii236.
- [38] I.M.L. Jöbses, H.C.H. Hiemstra, J.A. Roels, Fermentation kinetics of *Zymomonas mobilis* near zero growth rate, *Biotechnol. Bioeng.* 29 (4) (1987) 502–512.
- [39] P. Ottestad, Component analysis an alternative system, *Int. Stat. Rev.* 43 (1) (1975) 83–107.
- [40] I.T. Jolliffe, *Principal Component Analysis*, 2nd ed., Springer, New York, 2002.
- [41] J.E. Jackson, *A User's Guide to Principal Components*, John Wiley & Sons, Inc., New York, 1991.