

# Root cause analysis in multivariate statistical process monitoring: Integrating reconstruction-based multivariate contribution analysis with fuzzy-signed directed graphs

Bo He<sup>a,c</sup>, Tao Chen<sup>b,\*</sup>, Xianhui Yang<sup>a</sup>

<sup>a</sup> Department of Automation, Tsinghua University, Beijing 100084, China

<sup>b</sup> Department of Chemical and Process Engineering, University of Surrey, Guildford GU2 7XH, UK

<sup>c</sup> State Grid Energy Research Institute, Beijing 102209, China

\* Corresponding author; Tel.: +44 1483 686593; Email: t.chen@surrey.ac.uk

**Abstract** – Root cause analysis is an important method for fault diagnosis when used with multivariate statistical process monitoring (MSPM). Conventional contribution analysis in MSPM can only isolate the effects of the fault by pinpointing inconsistent variables, but not the underlying cause. By integrating reconstruction-based multivariate contribution analysis (RBMCA) with fuzzy-signed directed graph (SDG), this paper developed a hybrid fault diagnosis method to identify the root cause of the detected fault. First, a RBMCA-based fuzzy logic was proposed to represent the signs of the process variables. Then, the fuzzy logic was extended to examine the potential relationship from causes to effects in the form of the degree of truth (DoT). An efficient branch and bound algorithm was developed to search for the maximal DoT that explains the effect, and the corresponding causes can be identified. Except for the need to construct an SDG for the process, this method does not require historical data of known faults. The usefulness of the proposed method was demonstrated through a case study on the Tennessee Eastman benchmark problem.

**Keywords**- Fault diagnosis; Fuzzy logic; Multivariate statistical process monitoring; Reconstruction-based multivariate contributions; Signed directed graph.

## 1. Introduction

As competition becomes more and more intense in the process industry, coupled with the manufacturing plants getting more complex, the fault detection and diagnosis (FDD) technology has been recognised as an important tool to enable safe, efficient and environmentally benign operation of industrial processes (Frank, 1990; Venkatasubramanian et al., 2003a). Meanwhile, with the rapid progress in sensor technology, distributed process control and data acquisition systems, more and more process variables can be measured on a routine basis. As a result, data-driven FDD methods have attracted substantial attention in both academia and industry (Venkatasubramanian et al., 2003b; MacGregor & Cinar, 2012). Multivariate statistical process monitoring (MSPM) is a well-known data-driven FDD method (Qin, 2012; Yao & Gao, 2009). Traditional MSPM uses latent variable models, such as principal component analysis (PCA) and partial least squares (PLS), to detect faults and disturbances (Martin, Morris & Zhang, 1996; Qin, 2003). The extensions of PCA/PLS have enabled the monitoring of processes with more complex behaviours (e.g. non-Gaussian, non-linear and non-stationary); a sample of these methods may include kernel PCA (Choi et al., 2005), independent component analysis (Kano et al., 2003), PCA with one-class support vector machine (Ge & Song, 2011), dynamic PCA (Ku, Storer & Georgakis, 1995; Yao & Gao, 2007), among others. In practice, MSPM has been successfully implemented in many industrial applications (AlGhazzawi & Lennox, 2008; Kano & Nakagawa, 2008).

Following fault detection in MSPM are fault isolation and root cause analysis, which are important steps to help find the sources of the detected anomalies. To this end, if historical data are available for known types of fault, a wide range of data-drive methods can be used; for example the fault signatures (Yoon & MacGragor, 2001), fuzzy IF-THEN rules (Musulin, Yelamos & Puigjaner, 2006), statistical distances and angles (Raich & Cinar, 1997), fault subspaces (Dunia & Qin, 1998), fisher discriminant

analysis and support vector machine (Chiang, Kotanchek & Kordon, 2004). He et al. (2005) introduced a new method, in which discriminant analysis was used to obtain fault directions, and the directions are combined with contribution analysis for fault diagnosis. However, the faulty data are usually difficult to acquire during process operations; after all, abnormal process operation is a rare event by definition. Furthermore, these data-driven fault diagnosis methods may face difficulty in handling unknown faults, which have not occurred in the limited historical data.

An alternative approach is contribution analysis, which aims to isolate the variables that are the most responsible to the detected fault. Contribution analysis does not need prior knowledge of the faults, and thus is suitable for analysing unknown fault types (Miller, Swanson & Heckler, 1998). In addition, confidence limits can be established for contribution plots to further help discriminate the faulty variables from non-faulty ones (Westerhuis, Gurden & Smilde, 2000). However, the conventional contribution analysis is known to suffer from the fault “smearing” effect (Acala & Qin, 2009). To address this issue, the reconstruction-based (Acala & Qin, 2009) and missing variable (Chen & Sun, 2009) contribution methods have been proposed. In addition, the traditional contribution analysis focuses on the effect of individual variables, ignoring the interaction between variables. This limitation has recently been recognised, and the joint analysis of multiple variables under PCA models has been reported by using variable reconstruction (Liu, 2012), and missing variable analysis with a branch and bound (BnB) optimisation algorithm (Kariwala et al., 2010). We have followed this line of research to develop a generic, reconstruction-based multivariate contribution analysis (RBMCA) framework, which can be used with any process monitoring model (He et al., 2012). Later, this generic RBMCA approach has been further improved by introducing an  $L_1$ -penalty term for variable reconstruction (PRBMCA), prior to using BnB for optimisation (He et al., 2013).

Despite being effective in identifying the most responsible variables to the detected fault, contribution analysis still cannot directly reveal what fault has occurred and/or the root cause, which would be important information to help decide appropriate course of actions. Therefore, an approach to automatically interpret the variable contribution and to identify the root cause is needed to fill this gap. It appears that such methods would require the knowledge about the process being monitored, and the knowledge can be effectively represented in qualitative models.

Combing data-driven methods with qualitative modelling is not a new concept. Typical qualitative models include signed directed graph (SDG) (Vedam & Venkatasubramanian, 1999; Lee, Han & Yoon, 2004), expert system (Norvilas et al., 2000), cause-effect models (Leung & Romagnoli, 2002; Chiang & Braatz, 2003), plant connectivity modelling using extensible markup language (XML) (Thambirajah et al., 2009), among others. To handle the multivariate nature of process variables, the PCA-SDG hybrid (Vedam & Venkatasubramanian, 1999) is a well-known and effective method for fault detection and diagnosis. PCA-SDG relies on PCA to detect the fault, and then performs contribution analysis to determine the signs of the nodes (process variables) to be used in SDG. Then, a backward-forward propagation algorithm is used to search for the root cause. However, as discussed previously, the conventional contribution plot suffers from “fault smearing” (Acala & Qin, 2009) and faces difficulty in handling multivariate faults (He et al., 2012). In addition, the contribution threshold, which is used for determining the signs of nodes in SDG, is a tuning parameter and has strong impact on results. Furthermore, the crisp logic (as opposed to fuzzy logic) used in PCA-SDG, i.e. labelling the nodes as “positive”, “normal” or “negative”, does not reflect the quantitative status of a variable. For example, one node labelled “positive” may be closer to the “normal” situation than another “positive” node; however this cannot be modelled in crisp logic, resulting in poor resolution of diagnosis (Han, Shih & Lee, 1994).

In this paper, a root cause analysis method integrating PRBMCA with a fuzzy-SDG based reasoning scheme is proposed. PRBMCA determines the effect of the fault, in terms of the combinations of isolated faulty variables with corresponding signs; it inherits the advantage of MCA to avoid “fault smearing” and tuning the threshold for variable contribution. A reconstruction-based fuzzy logic is adopted to represent the signs of nodes in the form of degree of truth (DoT). The DoT function is further expanded to evaluate the “consistency” of “paths” between nodes, serving as the measure of likelihood of going

from the cause node to the effect in SDG. A depth-first search algorithm with BnB is developed to search for the paths from candidate cause nodes to effect nodes with respect to the DoT, and the candidate cause node with the maximal DoT of consistency is considered as the root cause. The use of fuzzy logic provides the quantitative information, in terms of DoT, of each cause. Moreover, in comparison with pure data-driven diagnosis methods, the proposed approach does not require historical data of known faults.

The hybrid fault diagnosis method also inherits another important feature of PRBMCA, i.e. it is a generic framework that is applicable to any process monitoring model, provided that a variable reconstruction algorithm can be developed for that particular model. Since the primary focus of this study is not fault detection but diagnosis, a simple probabilistic PCA (PPCA) is used as the process monitoring model (Kim & Lee, 2003; Chen & Sun, 2009) to demonstrate how the proposed method can help identify the root cause. Clearly, diagnosis is only possible for the faults that can be detected by the monitoring model. We will briefly discuss how other MSPM models can be incorporated in the proposed framework.

The rest of this paper is organised as follows. A brief review of RBMCA and PRBMCA is presented in Section 2. Section 3 presents the detailed the root cause analysis method, which combines PRBMCA and fuzzy-SDG. Section 4 discusses the results from the case study on the simulated Tennessee Eastman benchmark problem, and Section 5 concludes this paper.

## 2. Fault isolation using reconstruction-based multivariate contribution analysis

RBMCA is a general framework for isolating faulty variables and is applicable to any process monitoring model, provided that a variable reconstruction algorithm can be developed for that model (He et al., 2012). The basic idea of RBMCA is to search for the combination of process variables that contribute the most to the detected fault. A BnB algorithm was proposed to efficiently solve the combinatorial optimisation problem. Later, an  $L_1$  penalty was introduced to the variable reconstruction step of RBMCA (thus named PRBMCA), before using the BnB algorithm (He et al., 2013). PRBMCA tends to reduce the number of isolated faulty variables thus to give clearer interpretation of the results. The computation demand of PRBMCA is also significantly less than that of RBMCA, because the  $L_1$  penalty removes the improbable combinations of variables that would have to be evaluated in RBMCA. In this section, both versions are briefly reviewed.

### 2.1 Reconstruction-based multivariate contribution analysis

Suppose that a statistical model ( $\mathbf{M}$ ) has been developed to represent the normal process operations, and a monitoring statistic ( $D$ ) with appropriate control limit  $CL$  has been established to distinguish normal samples from faulty ones. We use  $D(\mathbf{x}|\mathbf{M})$  to denote the monitoring statistic of an  $n$ -dimensional sample  $\mathbf{x}$  under model  $\mathbf{M}$ . Thus, the process is considered out-of-control if  $D(\mathbf{x}|\mathbf{M}) > CL$ , and in-control if  $D(\mathbf{x}|\mathbf{M}) \leq CL$ . To enable variable reconstruction, a faulty sample  $\mathbf{x}$ , i.e.  $D(\mathbf{x}|\mathbf{M}) > CL$ , is divided into two parts:  $\mathbf{x} = \mathbf{x}^* + \mathbf{x}_f$ ,  $\mathbf{x}^*$  the fault-free part and  $\mathbf{x}_f = \mathbf{\Xi}_{\mathbf{x}_f} \cdot \mathbf{f}_{\mathbf{x}_f}$  the faulty part. Here,  $\mathbf{f}_{\mathbf{x}_f}$  is the magnitude of the fault and  $\mathbf{\Xi}_{\mathbf{x}_f}$  is an indicator matrix representing the combination of faulty variables (Qin, 2003). For instance, if the set of faulty variables is  $\mathbf{X}_f = \{x_2\}$ , then  $\mathbf{\Xi}_{\mathbf{x}_f} = [0 \ 1 \ 0 \ \dots]^T$ ; whilst if  $\mathbf{X}_f = \{x_2, x_4\}$ , the indicator matrix is

$$\mathbf{\Xi}_{\mathbf{x}_f} = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \end{bmatrix}^T.$$

Given such an indicator matrix, the magnitude of the fault can be determined by solving the following optimisation problem:

$$\hat{\mathbf{f}}_{\mathbf{x}_f} = \arg \min_{\mathbf{f}_{\mathbf{x}_f}} D(\mathbf{x} - \mathbf{\Xi}_{\mathbf{x}_f} \cdot \mathbf{f}_{\mathbf{x}_f} | \mathbf{M}) \quad (1)$$

This is equivalent to reconstructing the variables in the indicator matrix so that the reconstructed monitoring statistic reaches minimum (He et al., 2012). Then, the reconstruction-based multivariate contribution (RBMVC) of  $\mathbf{X}_f$  is:

$$\text{RBMVC}_{\mathbf{X}_f} = D(\mathbf{x} | \mathbf{M}) - D(\mathbf{x} - \mathbf{\Xi}_{\mathbf{x}_f} \cdot \hat{\mathbf{f}}_{\mathbf{x}_f} | \mathbf{M}) \quad (2)$$

where due to the minimisation in eq. (1),  $D(\mathbf{x}|\mathbf{M}) \geq D(\mathbf{x} - \mathbf{\Xi}_{\mathbf{x}_f} \cdot \hat{\mathbf{f}}_{\mathbf{x}_f}|\mathbf{M})$  and thus the contribution is non-negative.

The reconstruction in eq. (1) is a general formulation, and its solution depends on the choice of monitoring model. For the PPCA model used in this work,  $D(\mathbf{x}|\mathbf{M}) = \mathbf{x}^t \mathbf{C}^{-1} \mathbf{x}$ , where  $\mathbf{C}$  is the covariance matrix (Chen & Sun, 2009). Eq. (1) is thus an unconstrained quadratic optimisation problem which can be solved by using standard derivation-based algorithms. For the PPCA mixture model, an expectation-maximisation algorithm was developed to efficiently solve this problem in our previous study (He et al., 2012).

It should be noted that the indicator matrix  $\mathbf{\Xi}_{\mathbf{x}_f}$  does not need to be known *a priori*. In fact, the main task of RBMCA is to search for the set of variables, among  $2^n$  candidates, which should be reconstructed to give the best explanation of the fault. This is equivalent to finding  $\mathbf{\Xi}_{\mathbf{x}_f}$ , and it is a combinatorial optimisation problem. This problem has a trivial solution by reconstructing (isolating) all process variables, which motivated to include certain constraint on the number of variables that should be isolated (He et al., 2012). By isolating as fewer variables as possible (but still adequate to explain the fault), it also helps process engineers or root cause analysis algorithms to concentrate on the most influential variables. This concept was implemented by including a feasible solution set  $\mathbf{S}$  as follows (He et al., 2012):

$$\mathbf{S} = \{\mathbf{X}_f | \forall x_i \in \mathbf{X}_f, \Phi(\mathbf{X}_f \setminus x_i) \geq CL, \Phi(\mathbf{X}_f) < CL, \mathbf{X}_f \in \mathbf{X}_n\} \quad (3)$$

where  $\mathbf{X}_n = \{x_1, x_2, \dots, x_n\}$  is the set of all variables, and  $\Phi(\mathbf{X}_f) = D(\mathbf{x} - \mathbf{\Xi}_{\mathbf{x}_f} \cdot \hat{\mathbf{f}}_{\mathbf{x}_f}|\mathbf{M})$  is the monitoring statistic. The condition  $\Phi(\mathbf{X}_f) < CL$  ensures that the reconstruction brings the process back to normal, whilst  $\Phi(\mathbf{X}_f \setminus x_i) \geq CL$  indicates that by excluding any single variable, the reconstructed monitoring statistic is still out-of-control.

Subsequently, fault isolation was formulated as the following optimisation problem to find  $\mathbf{X}_S^{\text{RBMCA}}$ :

$$\Phi(\mathbf{X}_S^{\text{RBMCA}}) = \min_{\mathbf{X}_f \in \mathbf{S}} \Phi(\mathbf{X}_f) \quad (4)$$

which can be efficiently solved using a BnB algorithm (He et al., 2012).

## 2.2 Penalised reconstruction-based multivariate contribution analysis

In the above RBMCA method, a potential issue is that the magnitude of the fault,  $\mathbf{f}_{\mathbf{x}_f}$ , is not explicitly constrained. It was observed that RBMCA could isolate a large number of process variables when the fault magnitude is small or moderate, even with the constraint on the number of variables in eq. (3) (He et al., 2013). This is undesirable in practice, since it creates difficulty for the process engineers to find the underlying cause from many responsible variables. An alternative approach is to introduce a penalty on the magnitude of the fault, so that a set of potentially responsible variables are first isolated. The  $L_1$  penalty, which is well known to help retain only a few variables in multivariate optimisation problems (Tibshirani, 1996; Yi et al., 2011), is adopted. Subsequently, the candidate variables identified by the penalised reconstruction are further ‘‘refined’’ by using a BnB algorithm, similar to RBMCA. An overview of this two-step PRBMCA approach is given below.

**Step 1.** Use  $L_1$ -penalised reconstruction to obtain a candidate solution set  $\mathbf{X}_S^{L_1}$ . Specifically, the following optimisation problem is solved:

$$\min_{\hat{\mathbf{x}}_f} D(\mathbf{x} - \hat{\mathbf{x}}_f|\mathbf{M}), \quad \text{s. t. } |\hat{\mathbf{x}}_f|_{L_1} = \sum_{i=1}^n |\hat{x}_{f,i}| < t \quad (5)$$

The tuning parameter  $t$  is determined by

$$t_{\text{opt}} = \inf\{t | D(\mathbf{x} - \mathbf{x}_f(t)|\mathbf{M}) < CL\} \quad (6)$$

where  $\mathbf{x}_f(t) = \arg \min_{\hat{\mathbf{x}}_f} D(\mathbf{x} - \hat{\mathbf{x}}_f|\mathbf{M})$ , s.t.  $|\hat{\mathbf{x}}_f|_{L_1} = \sum_{i=1}^n |\hat{x}_{f,i}| < t$ . Standard binary search algorithm can be adopted to find  $t_{\text{opt}}$  efficiently. Then,  $\mathbf{X}_S^{L_1}$  consists of all the nonzero variables of  $\mathbf{x}_f(t_{\text{opt}})$ .

**Step 2.** Implement RBMCA, by using the BnB algorithm, to find the faulty variables from  $\mathbf{X}_S^{L_1}$ . This step guarantees that the final result belongs to the feasible solution set  $\mathbf{S}$ . More rigorously, the optimal solution  $\mathbf{X}_S^{\text{PRBMCA}}$  satisfies

$$\Phi(\mathbf{X}_S^{\text{PRBMCA}}) = \min_{\mathbf{X}_f \in \mathbf{S}} \Phi(\mathbf{X}_f), \quad \text{s. t. } \mathbf{X}_S^{\text{PRBMCA}} \subseteq \mathbf{X}_S^{L_1} \quad (7)$$

Similar to eq. (1), the reconstruction problem in eq. (5) is also formulated as a general framework, and the actual solution depends on the choice of monitoring model. When a PPCA model is used, eq. (5) can be solved by using quadratic programming (He et al., 2013). It was demonstrated that both the computational efficiency and the accuracy of fault isolation have been significantly improved by introducing the  $L_1$  penalty in RBMCA (He et al., 2013).

### 3. Fuzzy signed directed graphs for root cause analysis

Root cause analysis of detected fault is an important part of fault diagnosis. Signed directed graph (SDG) has been used to represent the cause-effect relations of process variables and to serve as the basis for root cause reasoning and identification (Iri et al., 1979; Han, Shih & Lee, 1994; Vedam & Venkatasubramanian, 1999; Lee, Han & Yoon, 2004; Lü et al., 2011). An SDG consists of nodes and arcs: nodes represent the variables while arcs represent cause-effect relations between nodes. The sign of a node represents the qualitative status of the variable, i.e. '+' (or '-') indicates the value of variable is higher (or lower) than normal, and '0' means the value is normal. The sign of arcs takes the value of '+' (or '-') to represent the cause and effect deviate in the same (or opposite) directions. The cause node of an arc is usually called initial node, while the effect one is called terminal node (Iri et al., 1979). As illustrated in the partial SDG of the Tennessee Eastman process in Fig. 1, the arc from  $x_1$  to  $x_{23}$  is '+', meaning that the former has a positive effect on the latter, i.e.  $x_{23}$  will be '+' if  $x_1$  is '+' (i.e.  $x_1$  has a positive deviation), and vice versa. Conversely,  $x_{30}$  will be '-' if  $x_{23}$  has a positive deviation since the arc from the latter to the former is '-'.

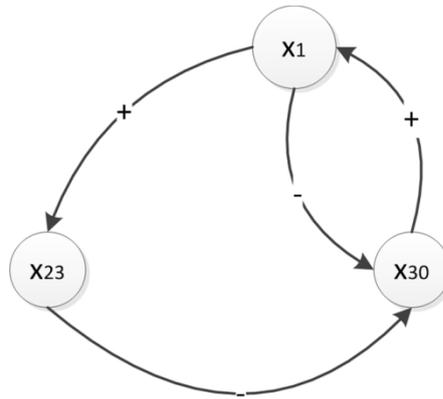


Fig 1. A partial SDG of the Tennessee Eastman process.

An SDG reflects the steady state gain around a certain operating point, and the signs may change when the process is non-linear and switched to a different operating point. As a result, multiple SDGs may need to be developed for each of the operating points. Fortunately, for a typical, well-operated plant, the number of operating points is usually small (e.g. a few). In addition, process engineers would have to understand the steady-state gains under each operating point anyway, in order to properly manage and operate the process. Translating this process knowledge into an SDG is extra work, but would be reasonable and justifiable given the added benefit of fault diagnosis. In turn, SDG would be a good way to document the steady-state process behaviour.

#### 3.1 Reconstruction-based fuzzy logic of the signs of nodes

Fuzzy logic is a form of many-valued logic. It has been successfully used in a variety of fields such as control theory (Lee, 1990), artificial intelligence (Dubois & Prade, 1991) and expert system (Zadeh, 1983). Compared with standard crisp logic in which the values are either 'true' or 'false', fuzzy logic evaluates the logic values in the form of the degree of truth (DoT) ranging from 0 to 1. For instance, a variable  $A$  can represent the reactor level with logic values of 'high' ('+') or 'low' ('-'), and  $\text{DoT}(A = +) = 0.9$ ,  $\text{DoT}(A = -) = 1 - \text{DoT}(A = +) = 0.1$ . It indicates that the level of reactor is 'fairly high'

and ‘slightly low’. As such, a central question in fuzzy logic is the design of membership functions that map a variable’s quantitative value (e.g. reactor level) to the DoT of logic values.

In the present work, each process variable (a node in the SDG) may take two logic values (signs): ‘+’ (positive deviation from its normal quantity) or ‘-’ (negative deviation from its normal quantity). The membership functions are defined as follows according to the fault isolation results from PRBMCA.

**Case 1:**  $x_i$  is isolated by PRBMCA as a faulty variable, and thus it is known to deviate from the normal value. If the original value,  $x_{ori,i}$ , is greater than the reconstructed one by PRBMCA,  $x_{rec,i}$ , then it is reasonable to say that the original value has a positive deviation with DoT of 1. Mathematically this is:

$$\text{DoT}(x_i = +) = \begin{cases} 1, & \text{if } x_{ori,i} > x_{rec,i} \\ 0, & \text{if } x_{ori,i} < x_{rec,i} \end{cases} \quad (8)$$

and  $\text{DoT}(x_i = -) = 1 - \text{DoT}(x_i = +)$ .

**Case 2:**  $x_i$  is not isolated by PRBMCA, and thus it should not be considered as definitely positive or negative. Instead, the DoT is designed according to how likely  $x_i$  deviates from its “normal value”, which, similar to Case 1, will be established through reconstruction. Recall that during the fault isolation stage, the reconstructed fault free sample (denoted by  $\mathbf{x}_{PRBMCA}$ ) does not change  $x_i$ , since  $x_i$  is not isolated by PRBMCA. Therefore, the following optimisation problem is solved to reconstruct the normal value of  $x_i$  (denoted by  $x_{rec,i}$ ):

$$x_{rec,i} = \underset{\hat{x}_{rec,i}}{\text{argmin}} D(\mathbf{x}_{PRBMCA} + \Xi_{\{x_i\}}(\hat{x}_{rec,i} - x_{ori,i}) | \mathbf{M}) \quad (9)$$

Subsequently, the DoT is defined as

$$\text{DoT}(x_i = +) = \begin{cases} 0.5 \left( 1 + \frac{D_{PRBMCA} - D_{rec,i}}{CL - D_{rec,i}} \right), & \text{if } x_{ori,i} > x_{rec,i} \\ 0.5 \left( 1 - \frac{D_{PRBMCA} - D_{rec,i}}{CL - D_{rec,i}} \right), & \text{if } x_{ori,i} \leq x_{rec,i} \end{cases} \quad (10)$$

and  $\text{DoT}(x_i = -) = 1 - \text{DoT}(x_i = +)$ . Here  $D_{PRBMCA}$  is the reconstructed monitoring statistic obtained by PRBMCA, and  $D_{rec,i}$  is the value of the objective function after solving the optimisation problem in eq. (9).

According to the PRBMCA framework and eq. (9),  $D_{rec,i} \leq D_{PRBMCA} \leq CL$ . Suppose  $D_{PRBMCA} - D_{rec,i} = 0$  (i.e. there is no difference in monitoring statistic by reconstructing  $x_i$ ), then  $\text{DoT}(x_i = +) = 0.5$ ,  $\text{DoT}(x_i = -) = 0.5$  (i.e. the value of  $x_i$  is neither too high nor too low). Conversely, the reconstruction of  $x_i$  results in significant reduction of monitoring statistic, if  $D_{PRBMCA} - D_{rec,i} \gg 0$ . It implies that the value of  $x_i$  is near totally high (if  $x_{rec,i} < x_{ori,i}$ ) or totally low (if  $x_{rec,i} > x_{ori,i}$ ). Thus,  $\text{DoT}(x_i = +) \approx 1$ ,  $\text{DoT}(x_i = -) \approx 0$  or  $\text{DoT}(x_i = +) \approx 0$ ,  $\text{DoT}(x_i = -) \approx 1$ . The denominator  $CL - D_{rec,i}$  normalizes the value from 0 to 1.

In summary, each of the process variables, no matter whether it has been isolated by PRBMCA, has the DoT of two logic values (‘+’ and ‘-’) according to eqs. (8)-(10).

### 3.2 DoT of path consistency

In SDG, reasoning is based on the connection from a cause node to an effect node. For a given signed effect node, a signed cause node is connected to it if  $\text{sign}(\text{effect node}) = \text{sign}(\text{cause node}) \cdot \prod_{i=1}^{\omega} \text{sign}(\text{arc}_i)$ , where  $\{\text{arc}_1, \text{arc}_2, \dots, \text{arc}_\omega\}$  is the sequence of arcs, or the path, from cause to effect. The path provides a potential explanation of the root cause of the process fault’s effect. In this work, the fuzzy logic is applied to evaluate the “consistency” of an arc that connects a cause node to an effect node:

$$\text{DoT}(\text{arc} = \text{consistent}) = \text{DoT}[\text{sign}(\text{effect node}) = \text{sign}(\text{cause node}) * \text{sign}(\text{arc})] \quad (11)$$

Moreover, the DoT of the consistency of a path is introduced to evaluate how trustworthy this explanation is:

$$\text{DoT}(\text{path} = \text{consistent}) = \text{DoT}(\text{sign}(\text{cause node})) \cdot \prod_{i=1}^{\omega} \text{DoT}(\text{arc}_i = \text{consistent}) \quad (12)$$

For instance, suppose that in Fig. 1, the effect is  $\{x_1, +\}$ , and a candidate path from the cause  $\{x_{23}, -\}$  is  $p = x_{23} \xrightarrow{-} x_{30} \xrightarrow{+} x_1$ , then

$$\text{DoT}(p = \text{consistent}) = \text{DoT}(x_{23} = -) \cdot \text{DoT}(\text{arc}_1 = \text{consistent}) \cdot \text{DoT}(\text{arc}_2 = \text{consistent})$$

$$= \text{DoT}(x_{23} = -) \cdot \text{DoT}(x_{30} = +) \cdot \text{DoT}(x_1 = +).$$

where  $\text{arc}_1 = x_{23} \xrightarrow{-} x_{30}$ ,  $\text{arc}_2 = x_{30} \xrightarrow{+} x_1$ .

### 3.3 DoT of cause to effect

In the context of SDG based fault diagnosis, root cause analysis aims to find the root node(s) that can explain the most effects, i.e. the root nodes have the most consistent paths to effects. Here, the effects are defined as the combinations of each faulty variable and sign isolated by PRBMCA, i.e. effect  $e = \{\text{var}_e, \text{sign}_e\}$ . For a given effect  $e$  and an assumed root cause  $c$  ( $c = \{\text{var}_c, \text{sign}_c\}$ ), there may exist multiple paths from  $c$  to  $e$ . By using the fuzzy logic method, the trustworthiness of each such path is evaluated in terms of the DoT of the consistency of the path. The path that gives the maximal DoT is taken to explain the effect, and more rigorously the maximal DoT is given by

$$\text{DoT}(\text{cause } c \text{ to effect } e) = \begin{cases} \max_{p \in \mathbf{P}} \text{DoT}(p), & \text{if } \mathbf{P} \neq \Phi \\ 0, & \text{if } \mathbf{P} = \Phi \end{cases} \quad (13)$$

where  $\mathbf{P}$  is the set of all paths from cause  $c$  to effect  $e$ .

Further, an identified faulty sample,  $\mathbf{x}_k$ , may have more than one process variable (thus effect) isolated by PRBMCA, and thus the overall DoT of the cause  $c$  is averaged over all effects as

$$\text{DoT}(\text{cause } c \text{ to } \mathbf{x}_k) = \frac{\sum_e \text{DoT}(\text{cause } c \text{ to effect } e)}{N_e} \quad (14)$$

where  $N_e$  is the number of effects determined by PRBMCA. Finally, for a sequence of faulty samples identified over a time period  $K$ ,  $\Delta = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K\}$

$$\text{DoT}(\text{cause } c \text{ to } \Delta) = \frac{\sum_{k=1}^K \text{DoT}(\text{cause } c \text{ to } \mathbf{x}_k)}{K} \quad (15)$$

The major challenges in this scheme are (i) to identify the candidate cause nodes, and (ii) to solve the combinatorial optimisation problem in eq. (13) for each candidate node. In the present work, the candidate root causes are simply obtained by enumerating all possible deviations of each variable. To address the second challenge, a depth-first search algorithm is proposed to find the optimal path from cause  $c$  to effect  $e$ . To further improve computational efficiency, the branch and bound (BnB) method is used. The key concept of BnB is to examine the lower bound of the objective function in eq. (13). If the upper bound of a set of candidate solutions is lower than the lower bound, then this entire set cannot be an optimal solution and thus can be pruned.

The search starts from the cause  $c$  towards the effect  $e$ . Each successor node of the current node is a sub-branch to divide the solution space into subsets. We use the tuple,  $\{\text{var}_i, \text{sign}_i, \text{DoT}_i\}$ , to represent the current node, where  $\text{DoT}_i$  is the DoT of the consistency of the path from  $c$  to the current node. The DoT of any successor of the current node may be expressed as  $\text{DoT}_i \cdot \text{DoT}(\text{path to the successor node})$ , which is no greater than  $\text{DoT}_i$  since by definition,  $\text{DoT} \in [0,1]$ . Therefore,  $\text{DoT}_i$  is the upper bound of all the sub-branches from the current node  $\{\text{var}_i, \text{sign}_i\}$ . In addition, we set the lower bound to be the maximal DoT of all paths that have been examined so far. The upper and lower bounds enabled the development of the BnB algorithm to solve the optimisation problem in eq. (13). The pseudo code of this BnB algorithm, by using a last-in-first-out stack as the data structure, is as follows.

```

DoT Search (cause  $c$ , SDG, effect  $e$ )

Set lower bound = 0;

Set stack=empty;

Stack_push({var $c$ , sign $c$ , DoT(var $c$  = sign $c$  )},stack);

while (stack!empty) do

    Stack_pop({var $top$ , sign $top$ , DoT $top$ }, stack);

    if (DoT $top$  ≥ lower bound) then

        if (var $top$  == var $e$ ) and (sign $top$  == sign $e$ ) then

            Set lower bound=DoT $top$ ;

        else if (var $top$ ! = var $e$ )

            foreach ( $suc$  is the successor of  $top$ ) and ( $suc$  is not in the set of predecessors of  $top$ )

                Stack_push({var $suc$ , sign $top$  · sign $arc$ , DoT $top$  · DoT(sign $suc$  = sign $top$  · sign $arc$ )}, stack);

            end

        end

    end

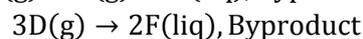
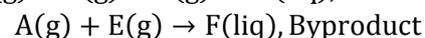
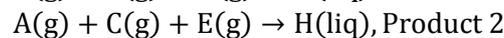
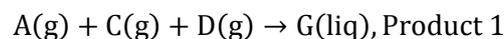
end

```

## 4. Case study

### 4.1 Process description

The Tennessee Eastman process (Downs & Vogel, 1993) is a well-known benchmark problem for testing process control FDD methods. There are five units in this plant: reactor, condenser, separator, stripper and compressor. This process produces two products (G and H) from four reactants (A, C, D, E). Meanwhile, an inert component B presents in stream 4 and a by-product F is produced as well. The reactions are:



The original process is open-loop and unstable. The reactor pressure grows quickly and exceeds the safe upper-limit, so that the plant must be shut down. In this paper, the plant-wide control structure developed by Lyman & Georgakis (1995) is adopted to maintain the process operation. The flow sheet of the process and the control strategy is shown in Fig. 2.

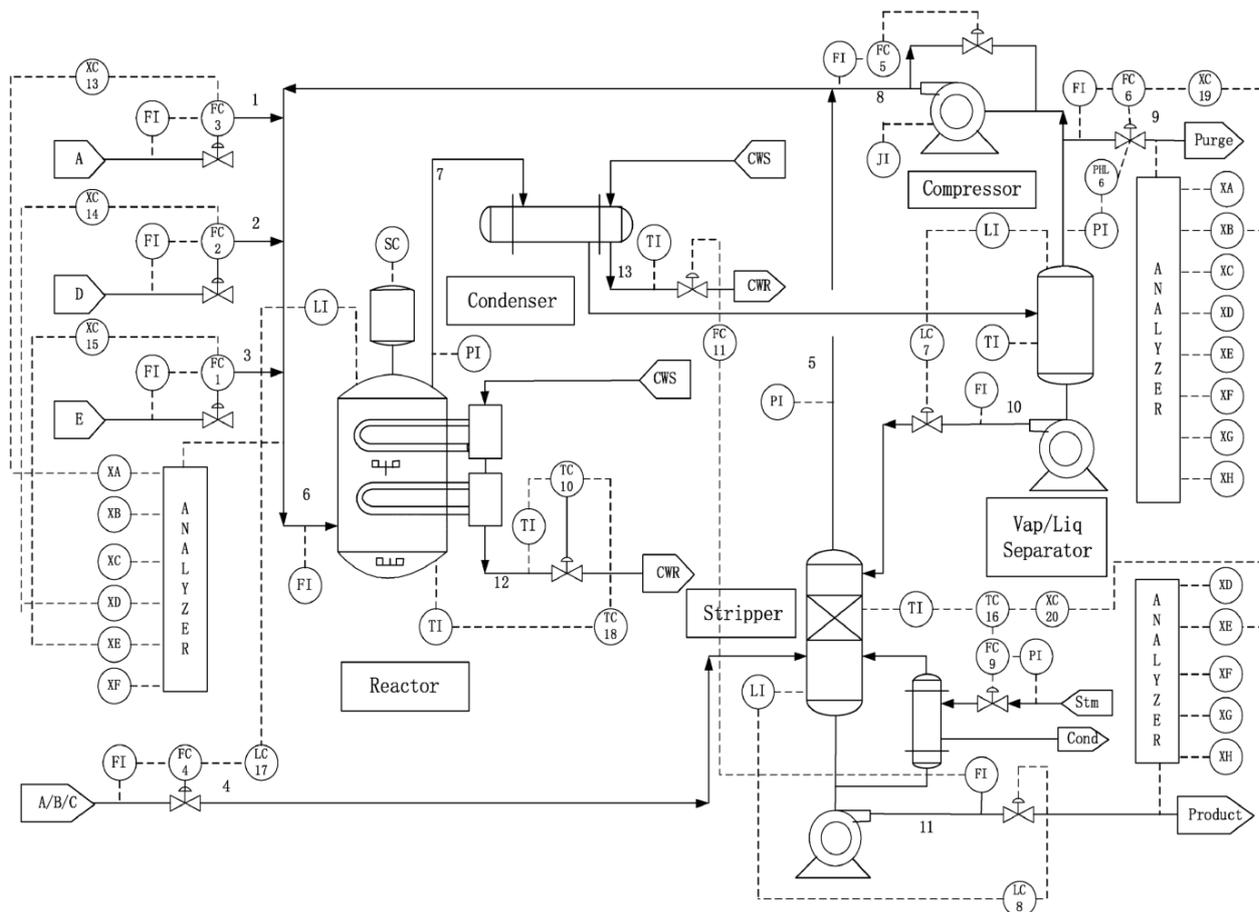


Fig. 2. Flow sheet of the Tennessee Eastman process.

For simplicity, only 38 process variables listed in Table 1 are selected in this work. Besides 22 continuous measured variables and 11 manipulated variables (MV), five composition measurements which presented in control loops are included. Out of the various process disturbances, Table 2 lists the seven typical faults that can be detected by PPCA (He et al., 2013), since PPCA has been chosen to illustrate the fault diagnosis results. As discussed previously, other statistical monitoring models can also be used within the PRBMCA-SDG framework, though a comprehensive exploration of alternative monitoring methods is outwith the scope of this study.

Table 1. Measured and manipulated variables.

ID	Description	ID	Description
$x_1$	A Feed(Stream1)	$x_{20}$	Compressor Work
$x_2$	D Feed(Stream2)	$x_{21}$	Reactor Cooling Water Outlet Temperature
$x_3$	E Feed(Stream3)	$x_{22}$	Condenser Cooling Water Outlet Temperature
$x_4$	A and C Feed(Stream4)	$x_{23}$	Component A (stream 6)
$x_5$	Recycle Flow(Stream 8)	$x_{24}$	Component D (stream 6)
$x_6$	Reactor Feed Rate(Stream6)	$x_{25}$	Component E (stream 6)
$x_7$	Reactor Pressure	$x_{26}$	Component B (stream 9)
$x_8$	Reactor Level	$x_{27}$	Component E (stream 11)
$x_9$	Reactor Temperature	$x_{28}$	MV to D feed flow (stream 2)
$x_{10}$	Purge Rate(Stream9)	$x_{29}$	MV to E feed flow (stream 3)
$x_{11}$	Product Separator Temperature	$x_{30}$	MV to A feed flow (stream 1)
$x_{12}$	Product Separator Level	$x_{31}$	MV to A and C feed flow (stream 4)
$x_{13}$	Product Separator Pressure	$x_{32}$	Compressor recycle valve
$x_{14}$	Product Separator Underflow(Stream 10)	$x_{33}$	Purge valve (stream 9)
$x_{15}$	Stripper Level	$x_{34}$	Separator pot liquid flow (stream 10)
$x_{16}$	Stripper Pressure	$x_{35}$	Stripper liquid product flow (stream 11)
$x_{17}$	Stripper Underflow(Stream 11)	$x_{36}$	Stripper steam valve
$x_{18}$	Stripper Temperature	$x_{37}$	MV to Reactor cooling water flow
$x_{19}$	Stripper Stream Flow	$x_{38}$	MV to Condenser cooling water flow

Table 2. Operational faults.

Fault ID	Description	Type
1	A/C feed ratio, B composition constant (stream 4)	Step
2	B composition, A/C- ratio constant (stream 4)	Step
4	Reactor cooling water inlet temperature	Step
5	Condenser cooling water inlet temperature	Step
6	A feed loss (stream 1)	Step
7	C header pressure loss-reduced availability (stream 4)	Step
11	Reactor cooling water inlet temperature	Random variation

## 4.2 SDG modelling

SDG modelling is the basis for reasoning and fault diagnosis in this study. An SDG model could be constructed by using quantitative knowledge, e.g. mathematical models (Maurya, Rengaswamy & Venkatasubramanian, 2003), or qualitative process knowledge or experiences (Vedam & Venkatasubramanian, 1999; Lee, Han & Yoon, 2004). In the present work, the SDG is constructed through qualitative analysis of the process flow sheet, since the accurate quantitative equations are not available. The key issue of building the SDG model of an industrial process (the TEP in this section) is to determine the signs of cause-effect relations of each pair of the variables. It should be noted that all the relations between variables shown in the SDG of this paper are founded upon the assumption of process running under the normal operation conditions. As such, these relations are valid only in the initial stage of the faulty operation, suggesting that only the early faulty samples should be used for diagnosis purpose. Arguably, the initial stage is the most critical time when the root-cause of the fault should be found, enabling timely mitigating measures to be taken.

The SDG model of TEP is presented in Table 3, indexed by the terminal variables. In this table, signed arcs between nodes in SDG are represented in the form of {initial variable, sign of arc} in the cells correspond to terminal variables. For example, for terminal variable  $x_1$ ,  $\{x_{30}, +\}$  means a positive arc from  $x_{30}$  to  $x_1$  (also illustrated in Fig. 1).

Table 3. SDG model of TEP

Terminal variables	Initial variables and sign of arcs	Terminal variables	initial variables and sign of arcs
$x_1$	$\{x_{30}, +\}$	$x_{20}$	$\{x_5, +\}, \{x_{10}, -\}, \{x_{13}, -\}$
$x_2$	$\{x_{28}, +\}$	$x_{21}$	$\{x_9, +\}, \{x_{37}, -\}$
$x_3$	$\{x_{29}, +\}$	$x_{22}$	$\{x_9, +\}, \{x_{11}, +\}, \{x_{38}, -\}$
$x_4$	$\{x_{31}, +\}$	$x_{23}$	$\{x_1, +\}, \{x_4, +\}, \{x_{24}, -\},$ $\{x_{25}, -\}$
$x_5$	$\{x_4, +\}, \{x_{10}, -\}, \{x_{13}, +\}, \{x_{20}, +$ $\{x_{32}, -\}$	$x_{24}$	$\{x_2, +\}, \{x_{23}, -\}, \{x_{25}, -\}$
$x_6$	$\{x_1, +\}, \{x_2, +\}, \{x_3, +\}$ $\{x_5, +\}, \{x_{16}, +\}$	$x_{25}$	$\{x_3, +\}, \{x_{23}, -\}, \{x_{24}, -\}$
$x_7$	$\{x_6, +\}, \{x_{23}, -\}$ $\{x_{24}, -\}, \{x_{25}, -\}, \{x_{26}, +\}$	$x_{26}$	$\{x_4, +\}$
$x_8$	$\{x_6, +\}, \{x_{11}, +\}$	$x_{27}$	$\{x_{18}, -\}$
$x_9$	$\{x_6, +\}, \{x_{23}, +\}, \{x_{24}, +\}$ $\{x_{25}, +\}, \{x_{32}, -\}$	$x_{28}$	$\{x_2, -\}, \{x_{24}, -\}$
$x_{10}$	$\{x_{33}, +\}$	$x_{29}$	$\{x_3, -\}, \{x_{25}, -\}$
$x_{11}$	$\{x_9, +\}, \{x_{38}, -\}$	$x_{30}$	$\{x_1, -\}, \{x_{23}, -\}$
$x_{12}$	$\{x_{11}, -\}, \{x_{34}, -\}$	$x_{31}$	$\{x_4, -\}, \{x_8, -\}$
$x_{13}$	$\{x_7, +\}, \{x_{11}, +\}$	$x_{32}$	$\{x_5, +\}$
$x_{14}$	$\{x_{12}, +\}, \{x_{34}, +\}$	$x_{33}$	$\{x_{10}, -\}, \{x_{26}, +\}$
$x_{15}$	$\{x_{35}, -\}, \{x_{36}, -\}$	$x_{34}$	$\{x_{12}, +\}$
$x_{16}$	$\{x_4, +\}, \{x_{18}, +\}$	$x_{35}$	$\{x_{15}, +\}$
$x_{17}$	$\{x_{15}, +\}, \{x_{35}, +\}$	$x_{36}$	$\{x_{18}, -\}, \{x_{19}, -\}, \{x_{27}, +\}$
$x_{18}$	$\{x_4, -\}, \{x_{14}, +\}, \{x_{19}, +\}$	$x_{37}$	$\{x_9, +\}, \{x_{21}, +\}$
$x_{19}$	$\{x_{36}, +\}$	$x_{38}$	$\{x_{17}, -\}$

Confined by the length of this article, five typical cases are selected in this section to illustrate how to build this SDG. Other the relations are obtained in a similarly way.

Case 1: control loop around inlet stream 1. As shown in Fig. 2, a cascade PID control loop is used to regulate the mole-percentage of component A ( $x_{23}$ ) (the main loop output) by manipulating the flow rate of stream 1 ( $x_1$ ), which is directly controlled in the inner loop by its MV ( $x_{30}$ ). For the main loop, if  $x_{23}$  increases, feeding too much reactant A to the reactor, the MV will decrease to reduce the flow rate of stream 1. Similarly in the inner loop, if the flow rate of stream 1 ( $x_1$ ) is higher than normal, the MV will also decrease. Consequently,  $x_1$  and  $x_{23}$  have negative effects on  $x_{30}$  (c.f. terminal variable  $x_{30}$  in Table 3). Clearly, the MV of stream 1 ( $x_{30}$ ) has a positive effect on the flow rate of stream 1 ( $x_1$ ). Furthermore, increasing  $x_1$  will increase the mole fraction of component A in stream 6; hence  $x_1$  has a positive effect on  $x_{23}$ . These cause-effect relations are also illustrated in Fig. 1.

Case 2: heat exchange around the reactor. Because of the heat exchanger, the reactor temperature ( $x_9$ ) has a positive effect on the outlet temperature of the cooling water ( $x_{21}$ ). In addition,  $x_{37}$  manipulates the flow of the reactor cooling water, and it has a negative effect on the cooling water outlet temperature ( $x_{21}$ ).

Case 3: mass balance. Because the reactor feed rate ( $x_6$ ) is the sum of  $x_1, x_2, x_3, x_5$ , each flow has a positive effect on it. Similarly,  $x_6$  will increase if there is more gas out from the stripper, i.e. the pressure of stripper ( $x_{16}$ ) has a positive effect on  $x_6$ . In addition, more (less) reactor feed results in higher (lower) reactor pressure; hence  $x_6$  has a positive effect on  $x_7$  (the reactor pressure). Further, the pressure of separator ( $x_{13}$ ) is positively affected by  $x_7$  due to the gas propagation in the plant.

Case 4: reaction. The reactions produce more products and release more heat (the reactions are exothermic), if the feed rate is higher (provided that the stoichiometry is within a proper range). Thus, the feed rate  $x_6$  has a positive effect on the reactor level ( $x_8$ ) and temperature ( $x_9$ ). Moreover, if any of the mole fraction of reactants in stream 6 decreases, the reaction slows down with excessive unreacted gas, causing increase in the reactor pressure, and decrease in the reactor level and temperature. Therefore,  $x_{23}$ ,  $x_{24}$  and  $x_{25}$  have negative effect on  $x_7$ , and positive effect on  $x_8$  and  $x_9$ .

Case 5: vaporisation in the separator and stripper. Clearly, the pressure in the separator ( $x_{13}$ ) is positively affected by the separator temperature ( $x_{11}$ ). Moreover, the liquid component E ( $x_{27}$ ) in the product flow is negatively affected by the temperature of the stripper ( $x_{18}$ ).

### 4.3 Results and analysis

The simulation data are generated by Russell, Chiang & Braatz (2000) and can be downloaded from <http://web.mit.edu/braatzgroup/links.html>. There are 500 samples under normal operation, which are used to build a PPCA model (Kim & Lee, 2004; Chen & Sun, 2009) with 14 principal components, which explained 75.98% variance. Following (Chen & Sun, 2009), the single likelihood-based control limit is set to be the 99%-fractile of the  $\chi^2$  distribution with degrees of freedom of  $n=38$  (the number of process variables), i.e.  $CL = \chi_{38}^2(0.99) = 61.16$ . For each faulty dataset, 960 samples are available whereby the abnormal events are triggered after 160 samples. The sample interval of the continuous measured variables ( $x_1$  to  $x_{22}$ ) and MVs ( $x_{28}$  to  $x_{38}$ ) is 3 min, whilst that of  $x_{23}$  to  $x_{26}$  is 6 min and of  $x_{27}$  is 15 min. For the purpose of modelling, the sample interval is kept at 3 min; zero-order holding is adopted for the analyser measurements ( $x_{23}$  to  $x_{27}$ ) with longer intervals.

Once a fault is detected by PPCA, PRBMCA is used to isolate the faulty variables. As illustration, the set of isolated variables from the 6th sample of fault 7 is:  $X_f = \{x_2, x_4, x_6, x_7, x_{11}, x_{13}, x_{16}, x_{18}, x_{20}, x_{21}, x_{31}, x_{32}, x_{37}\}$ , and the reconstructed monitoring statistic is 49.50. The set,  $X_f$ , happens to be the only solution of the PRBMCA algorithm. If any variable in  $X_f$  is excluded, the reconstructed monitoring statistic is greater than the control limit (61.16). On the other side, including any additional variable is unnecessary according to the definition of the feasible region in eq. (3), because the reconstructed statistic is already below the control limit. However, the optimisation problem could have multiple solutions with similar objective values, and thus it does not guarantee to find the actual root cause or faulty variables. Since the present study is focused on the integration of fault isolation with SDG for fault diagnosis, the reader is referred to more detailed investigation of the fault isolation accuracy in (He et al., 2013).

It appears that fault 7 has complex impact on a range of process variables, and it is very difficult to pinpoint the underlying cause of the fault by simply examining these isolated variables. Next, the DoT of the signs of these isolated variables (effect nodes) is calculated according to eqs. (8)-(10), and illustrated in Fig. 3. For comparison, the DoT of the remaining “normal” variables is also displayed. It can be observed that for identified effect nodes, the DoT strongly favours either a positive (“+”) or a negative (“-”) deviation from normal situation. In contrast, the DoT of the normal variables tends to be closer to 0.5, since they are not isolated by PRBMCA.

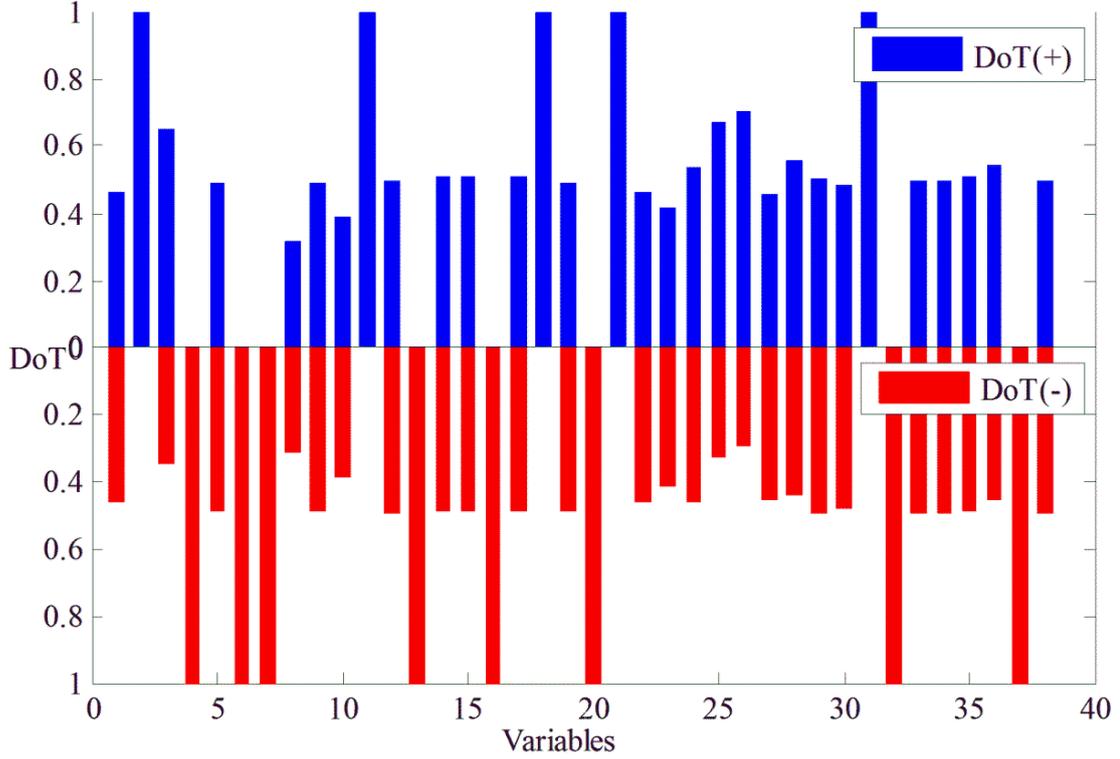


Fig. 3: DoT of the signs of the process variables, corresponding to the 6th sample of fault 7.

After obtaining the DoT of the effect nodes, the next step is to find the most probable cause node(s). The DoT values of all possible cause nodes (76 in total: 38 variables with “+” or “-” deviation) is calculated. Table 4 shows the five most probable causes and the corresponding DoT values for fault 7. The DoT of  $\{x_4, -\}$  (reduction of inlet feed of stream 4) is 0.71, which is significantly greater than any other cause. It suggests that  $\{x_4, -\}$  is the root cause of fault 7.

Table 4. The five most probable causes for fault 7.

Causes	$\{x_4, -\}$	$\{x_6, -\}$	$\{x_{20}, -\}$	$\{x_5, -\}$	$\{x_2, +\}$
DoT	0.71	0.41	0.33	0.29	0.25

Further, Table 5 shows the path and DoT from the root cause  $\{x_4, -\}$  to the effects for fault 7. All the effects can be explained by the root cause  $\{x_4, -\}$  with a sequence of deviations of variables, represented as paths in this table. It should be noted that along a path, not all the deviations are isolated by PRBMCA, which is why fuzzy logic, not crisp logic, is desired. For instance, by using crisp logic, the effect  $\{x_{32}, -\}$  could not be explained by the cause  $\{x_4, -\}$ , because  $x_5$  is not isolated and crisp logic would give  $\text{DoT}[\{x_5, -\}] = 0$ , meaning that the path  $\{x_4, -\} \rightarrow \{x_5, -\} \rightarrow \{x_{32}, -\}$  is not possible. However, Table 5 shows that  $\{x_4, -\}$  could be considered as the cause of the effect  $\{x_{32}, -\}$ , if  $x_5$  is lower than normal; the DoT of the path (0.49) gives the trustworthiness of this explanation. The DoT is also used to compare competing explanations. For example, the effect  $\{x_{32}, -\}$  can also be explained by the cause  $\{x_1, -\}$  through the path:  $\{x_1, -\} \rightarrow \{x_6, -\} \rightarrow \{x_7, -\} \rightarrow \{x_{13}, -\} \rightarrow \{x_5, -\} \rightarrow \{x_{32}, -\}$ , whose DoT, however, is 0.23 and less than that of  $\{x_4, -\}$ . It suggests that for the effect  $\{x_{32}, -\}$ ,  $\{x_4, -\}$  is a more probable cause than  $\{x_1, -\}$ . In summary, fuzzy logic provides quantity information of every candidate cause in the form of DoT, thus better analysis of the potential root cause than binary crisp logic does.

Table 5. Paths and DoTs from root cause  $\{x_4, -\}$  to effects of the 6th sample of fault 7.

Effects	Path	DoT
$\{x_2, +\}$	$\{x_4, -\} \rightarrow \{x_{23}, -\} \rightarrow \{x_{25}, +\} \rightarrow \{x_{24}, -\} \rightarrow \{x_{28}, +\} \rightarrow \{x_2, +\}$	0.07
$\{x_4, -\}$	$\{x_4, -\}$	1
$\{x_6, -\}$	$\{x_4, -\} \rightarrow \{x_{16}, -\} \rightarrow \{x_6, -\}$	1
$\{x_7, -\}$	$\{x_4, -\} \rightarrow \{x_{16}, -\} \rightarrow \{x_6, -\} \rightarrow \{x_7, -\}$	1
$\{x_{11}, +\}$	$\{x_4, -\} \rightarrow \{x_5, -\} \rightarrow \{x_{32}, -\} \rightarrow \{x_9, +\} \rightarrow \{x_{11}, +\}$	0.24
$\{x_{13}, -\}$	$\{x_4, -\} \rightarrow \{x_{16}, -\} \rightarrow \{x_6, -\} \rightarrow \{x_7, -\} \rightarrow \{x_{13}, -\}$	1
$\{x_{16}, -\}$	$\{x_4, -\} \rightarrow \{x_{16}, -\}$	1
$\{x_{18}, +\}$	$\{x_4, -\} \rightarrow \{x_{18}, +\}$	1
$\{x_{20}, -\}$	$\{x_4, -\} \rightarrow \{x_5, -\} \rightarrow \{x_{20}, -\}$	0.49
$\{x_{21}, +\}$	$\{x_4, -\} \rightarrow \{x_{16}, -\} \rightarrow \{x_6, -\} \rightarrow \{x_9, -\} \rightarrow \{x_{37}, -\} \rightarrow \{x_{21}, +\}$	0.49
$\{x_{31}, +\}$	$\{x_4, -\} \rightarrow \{x_{31}, +\}$	1
$\{x_{32}, -\}$	$\{x_4, -\} \rightarrow \{x_5, -\} \rightarrow \{x_{32}, -\}$	0.49
$\{x_{37}, -\}$	$\{x_4, -\} \rightarrow \{x_{16}, -\} \rightarrow \{x_6, -\} \rightarrow \{x_9, -\} \rightarrow \{x_{37}, -\}$	0.49

As comparison, the univariate reconstruction-based contribution plot for the 6th sample of fault 7 is given in Fig. 4. Notice that the contributions and control limits are all shown in logarithmic scale for ease of illustration, i.e. the univariate contributions are  $\log[\text{RBMVC}_{x_i}]$ ,  $i = 1, 2 \dots n$ , and the control limit is  $\log[D(\mathbf{x}|\mathbf{M}) - \text{CL}]$ . Fig. 4 shows that no variable's contribution exceeds the control limit, suggesting that the univariate analysis does not unambiguously isolate the faulty variables, and thus is less helpful (if at all in this case) for root cause analysis. Besides, variable  $x_{16}$  (stripper pressure) with the largest univariate contribution may be incorrectly identified as the root cause of this fault.

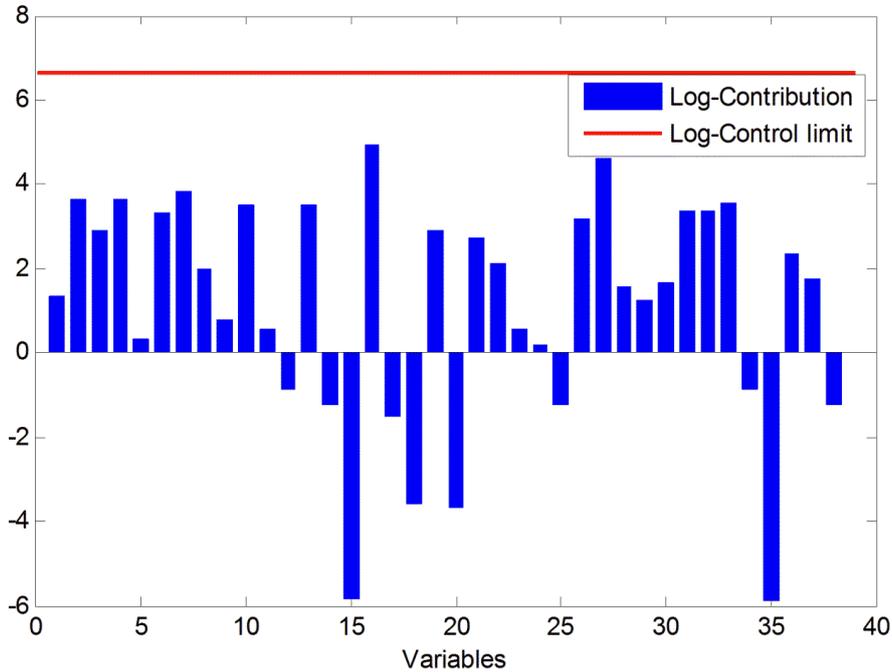


Fig. 4. Univariate reconstruction-based contribution for fault 7.

Table 6 summarises the results for all fault types. For each fault, the root cause analysis is carried out every 15 min (the longest sampling interval of process variables) until 1 hr, after the fault is detected. At a particular time point, all faulty samples up to that time are used for root cause analysis according to eq. (15). The cause node with the highest DoT is determined as the root cause.

Table 6. Summary of the RCA results.

Fault ID	0 min	DoT	15 min	DoT	30 min	DoT	45 min	DoT	60 min	DoT
1	{ $x_{21}, +$ }	1.00	{ $x_7, +$ }	0.58	{ $x_7, +$ }	0.65	{ $x_{23}, -$ }	0.64	{ $x_{23}, -$ }	0.67
2	{ $x_{26}, +$ }	1.00	{ $x_{26}, +$ }	0.85	{ $x_{26}, +$ }	0.85	{ $x_{26}, +$ }	0.84	{ $x_{26}, +$ }	0.83
4	{ $x_9, +$ }	1.00	{ $x_{37}, +$ }	0.90	{ $x_{37}, +$ }	0.95	{ $x_{37}, +$ }	0.84	{ $x_{37}, +$ }	0.86
5	{ $x_{22}, +$ }	1.00	{ $x_{11}, +$ }	0.70	{ $x_{11}, +$ }	0.71	{ $x_{11}, +$ }	0.66	{ $x_{11}, +$ }	0.65
6	{ $x_1, -$ }	1.00	{ $x_1, -$ }	0.76	{ $x_1, -$ }	0.61	{ $x_1, -$ }	0.53	{ $x_1, -$ }	0.52
7	{ $x_4, -$ }	1.00	{ $x_4, -$ }	0.78	{ $x_4, -$ }	0.76	{ $x_4, -$ }	0.75	{ $x_4, -$ }	0.64
11	{ $x_{37}, -$ }	1.00	{ $x_{37}, -$ }	0.70	{ $x_9, +$ }	0.48	{ $x_9, -$ }	0.35	{ $x_9, -$ }	0.37

The root cause of fault 1, due to a step change in A/C feed ratio in stream 4, cannot be directly attained from SDG, because the compositions of component A and C in stream 4 are not measured. However, the most probable causes are identified as the reduction of mole-percentage of component A in stream 6 ( $\{x_{23}, -\}$ ) and the increasing pressure of reactor ( $\{x_7, +\}$ ). Note that stream 4 goes through the stripper, and the gases are recycled to join streams 1, 2 and 3 to form stream 6, causing reduced A and increased reactor pressure (c.f. discussion of Case 4 in Section 4.2). Although  $\{x_{21}, +\}$  is identified as the root cause for the 0 min after fault detected, it is excluded since its DoT is very small (not shown in the table) in the rest of 60 minutes. Although  $\{x_{23}, -\}$  and  $\{x_7, +\}$  are still only the phenomena of the fault, they can help the engineer narrow down the possibilities. Furthermore, the only other source of component A in stream 6 is stream 1, which can be ruled out as the cause since the DoT of  $\{x_1, +\}$  is very low (not shown in the table). Therefore, the engineer can determine the root cause as the reduction of component A in stream 4, with the help of the proposed method and careful analysis.

As for fault 2, the root cause is too much component B in stream 4. Accordingly,  $\{x_{26}, +\}$  is identified with high DoT, indicating too much component B in the purge. Since stream 4 is the only source of this component, it should be examined.

Faults 4 and 5 are due to step change in cooling water temperature, which are not directly measured. For fault 4, the results show that the root cause could be the rise in the MV of the reactor cooling water ( $\{x_{37}, +\}$ ) or temperature rise in the reactor ( $\{x_9, +\}$ ). Further investigation could be carried out to ascertain whether  $\{x_{37}, +\}$  and/or  $\{x_9, +\}$  is the underlying cause, or if they are ruled out it is the inlet cooling water that is causing the problem. Similarly for fault 5, the root cause obtained by the proposed method is the high separator temperature ( $\{x_{11}, +\}$ ) or the outlet cooling water of the condenser ( $\{x_{22}, +\}$ ), which is useful information for further analysis, including the decision to check the cooling water inlet temperature.

In the case of fault 6, all results indicate that the root cause is the reduction of flow of stream 1 ( $\{x_1, -\}$ ). Similarly, all results for fault 7 point to the reduction of flow of stream 4 ( $\{x_4, -\}$ ). In both cases, the identified causes are helpful for supporting further investigation.

Unlike the previous step disturbances, fault 11 is due to random variation of reactor cooling water inlet temperature. The most probable causes are the variation of reactor temperature ( $\{x_9, +\}, \{x_9, -\}$ ) and MV ( $\{x_{37}, -\}$ ). Although the DoT of  $\{x_9, -\}$  is only 0.35 and 0.37 after 30 min, that of  $\{x_9, +\}$  is also significant (0.31 for 45 min and 0.32 for 60 min), which pointing to the variation of  $x_9$  to both directions.

The results show that DoT in general decreases over time. This is because that after fault occurs, the process deviates from the steady state, from which the SDG was built. As a result, the proposed method should be used with extra care when including a long period of data after the fault is detected. Nevertheless, the results suggest that the method still provides practically useful information to the operator/engineer. For example for fault 7, the DoT of  $x_4$  drops from 1 to 0.64; but the decrease in  $x_4$  should still be considered as the root cause, since  $\{x_4, -\}$  always has the highest DoT from 0 to 60 min.

Table 7 reports the computation time of the proposed method, obtained under MATLAB<sup>®</sup> 2011a on a desktop computer with Intel<sup>®</sup> Core<sup>™</sup> Duo CPU E8300 (2.83 GHz). The computation appears fairly reasonable in most cases, except that it takes over 15 min (meaning it is longer than the longest sampling

interval) at 60 min after fault 7 has been detected. It should be noted that 38 variables are used in the case study; for systems with even more process variables, the computation will further increase, and even the PRBMCA and BnB methods may not be able to solve the combinatorial optimisation problem in good time. In addition, for very large scale, complex systems, the development of an SDG can also pose an issue. Therefore, the proposed method in its current form may be limited to medium scale processes (though there are many such systems), and future research is needed to further improve the computational efficiency. Related discussions of computation in the fault isolation step can be found in (He et al., 2012, 2013; Kariwala et al., 2010).

Table 7. Summary of the computation time in seconds.

Fault ID	0 min	15 min	30 min	45 min	60 min
1	0.3	4.9	40.9	132.1	242.4
2	1.3	17.3	43.1	81.9	104.9
4	0.3	1.4	2.8	16.9	18.9
5	0.3	10.6	23.2	74.9	151.9
6	8.5	64.5	157.8	300.9	414.2
7	10.4	114.6	322.4	665.3	1030.5
11	0.6	1.5	6.4	8.1	11.6

## 5. Concluding remarks

This paper reports a root cause analysis method for MSPM, integrating reconstruction-based multivariate contribution analysis with fuzzy-signed directed graphs. It fills the gap between fault isolation through multivariate contribution analysis, and the need to drill down to the underlying cause of the fault. Except for the need to develop an SDG that represents the cause-effect relationship between process variables, this method does not need other prior knowledge or historical dataset of known process faults in the entire procedure of detection, isolation and root cause analysis. We acknowledge that the SDG does not always reveal the exact root cause, in particular when the cause related variables are not directly measured. Nevertheless, the method is effective in pointing to the possible explanations, which may need to undergo further investigation. On the other hand, we argue that in practice, fully automatic fault detection, isolation and diagnosis without human intervention may not be possible or desired. Instead, the provision of a decision support tool that incorporates process knowledge may be more welcome. The case study through the Tennessee Eastman benchmark problem shows that the method can provide useful information to support the diagnosis of detected abnormal process operations.

In terms of computation, the proposed method may be limited to processes with tens of variables (38 in the case study), and future research is needed to further improve the computational efficiency when solving the combinatorial optimisation problem.

## Acknowledgement

This work was carried out during Bo He's visit to the University of Surrey, supported by the Santander Universities through a Santander Postgraduate Research Award.

## References

- Alcala, C. F., Qin, S. J. (2009). Reconstruction-based contribution for process monitoring. *Automatica*, 45, 1593-1600.
- Alcala, C. F., Qin, S. J. (2011). Analysis and generalization of fault diagnosis methods for process monitoring. *Journal of Process Control*, 21, 322-330.
- AlGhazzawi, A., Lennox, B. (2008). Monitoring a complex refining process using multivariate statistics. *Control Engineering Practice*, 16, 294-307.
- Chen, T., Sun, Y. (2009). Probabilistic contribution analysis for statistical process monitoring: A missing variable approach. *Control Engineering Practice*, 17, 469-477.
- Chiang, L.H., Braatz, R.D. (2003). Process monitoring using cause map and multivariate statistics: fault detection and identification. *Chemometrics and Intelligent Laboratory Systems*, 65, 159-178.
- Chiang, L.H., Kotanchek, M.E., Kordon, A.K. (2004). Fault diagnosis based on fisher discriminant analysis and support vector machines, *Computers and Chemical Engineering*, 28, 1389-1401.
- Choi, S.W., Lee, C., Lee, J., Park, J.H., Lee, I. (2005). Fault detection and identification of nonlinear processes based on kernel PCA. *Chemometrics and Intelligent Laboratory Systems*, 75, 55-67.
- Downs, J. J., Vogel, E. F. (1993). A plant-wide industrial process control problem, *Computers and Chemical Engineering*, 17, 245-255.
- Dubois, D., Prade, H. (1991). Fuzzy sets in approximate reasoning, Part 1: Inference with possibility distributions, *Fuzzy Sets and Systems*, 40, 143-202
- Dunia, R., Qin, S.J. (1998). Subspace approach to multidimensional fault identification and reconstruction. *AIChE Journal*, 44, 1813-1831.
- Frank, P.M. (1990) Fault diagnosis in dynamic system using analytical and knowledge-based redundancy-A survey and some new results. *Automatic*, 26, 459-474.
- Ge, Z., Song, Z. (2011). A distribution-free method for process monitoring. *Expert Systems with Applications*, 38, 9821-9829.
- Han, C., Shih, R., Lee, L. (1994). Quantifying signed directed graphs with the fuzzy set for fault diagnosis resolution improvement. *Industrial and Engineering Chemistry Research*, 33, 1943-1954.
- He, B., Yang, X., Chen, T., Zhang, J. (2012). Reconstruction-based multivariate contribution analysis for fault isolation: A branch and bound approach. *Journal of Process Control*, 22, 1228-1236.
- He, B., Zhang, J., Chen, T., Yang, X. (2013). Penalized reconstruction-based multivariate contribution analysis for fault isolation. *Industrial & Engineering Chemistry Research*, 52, 7784-7794.
- He, Q.P., Qin, S.J., Wang, J. (2005). A new fault diagnosis method using fault directions in Fisher discriminant analysis, *AIChE Journal*, 51, 555-571.
- Iri, M., Aoki, K., O'Shima, E., Matsuyama, H. (1979). An algorithm for diagnosis of system failures in the chemical process. *Computers and Chemical Engineering*, 3, 489-493.
- Kano, M., Nakagawa, Y. (2008). Data-based process monitoring, process control and quality improvement: Recent developments and applications in steel industry. *Computers and Chemical Engineering*, 32, 12-24.
- Kano, M., Tanaka, S., Hasebe, S., Hashimoto, I. (2003). Monitoring Independent Components for fault detection. *AIChE Journal*, 49, 969-976.
- Kariwala, V., Odiwei, P. E., Cao, Y., Chen, T. (2010). A branch and bound method for isolation of faulty variables through missing variable analysis, *Journal of Process Control*, 20, 1198-1206.
- Kim, D., Lee, I.-B. (2003). Process monitoring based on probabilistic PCA. *Chemometrics and Intelligent Laboratory Systems*, 67, 109-123.
- Ku, W., Storer, R.H., Georgakis, C. (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 30, 179-196.
- Lee, C. (1990). Fuzzy logic in control systems: fuzzy logic controller-part I. *IEEE Transactions on Systems, Man, and Cybernetics*, 20, 404-418.

- Lee, G., Han, C., Yoon, E. (2004). Multiple-Fault diagnosis of the Tennessee Eastman process based on system decomposition and dynamic PLS. *Industrial and Engineering Chemistry Research*, 43, 8037-8048.
- Leung, D., Romagnoli, J. (2002) . An integration mechanism for multivariate knowledge-based fault diagnosis. *Journal of Process Control*, 12, 15-26.
- Liu, J. (2012). Fault diagnosis using contribution plots without smearing effect on non-faulty variables. *Journal of Process Control*, 22, 1609-1623.
- Lü, N., Xiong, Z., Wang, X., Ren, C. (2011). Integrated framework of probabilistic signed digraph based fault diagnosis approach to a gas fractionation unit. *Industrial and Engineering Chemistry Research*, 50, 10062-10073.
- Lyman, P.R., Georgakis, C. (1995). Plant-wide control of the Tennessee Eastman problem. *Computers and Chemical Engineering* , 19, 321-331.
- MacGregor, J., Cinar, A. (2012). Monitoring, fault diagnosis, fault-tolerant control and optimization: Data driven method. *Computers and Chemical Engineering*, 47, 111-120.
- Martin, E. B., Morris, A. J., Zhang, J. (1996). Process performance monitoring using multivariate statistical process control, *IEE Proceedings: Control Theory and Applications*, 143, 134-144
- Maurya, M.R., Rengaswamy, R., & Venkatasubramanian,V. (2003) A Systematic Framework for the Development and Analysis of Signed Digraphs for Chemical Processes.1. Algorithms and analysis. *Industrial and Engineering Chemistry Research*, 42, 4789-4810.
- Miller, P., Swanson, R.E., Heckler, C. F. (1998). Contribution plots: A missing link in multivariate quality control. *International Journal of Applied Mathematics and Computer Science*. 8, 775-792.
- Musulini, E., Yelamos, I., Puigjaner, L. (2006). Integration of principal component analysis and fuzzy logic system for comprehensive process fault detection and diagnosis. *Industrial and Engineering Chemistry Research*, 50, 1739-1750.
- Norvilas, A., Negiz, A., DeCicco, J., Cinar, A. (2000). Intelligent process monitoring by interfacing knowledge knowledge-based systems and multivariate statistical monitoring. *Journal of Process Control*, 10, 341-350.
- Qin, S.J. (2003).Statistical process monitoring: basics and beyond. *Journal of Chemometrics*, 17, 480-502.
- Qin, S.J. (2012).Survey on data-driven industrial process monitoring and diagnosis. *Annual Reviews in Control*, 36, 220-234
- Raich, A. Cinar, A. (1997). Diagnosis of process disturbances by statistical distance and angle measure. *Computers and Chemical Engineering* , 6, 661-673
- Russell, E.L., Chiang, L.H., Braatz, R.D. (2000). Data-driven techniques for fault detection and diagnosis in chemical processes. London:Springer-Verlag.
- Thambirajah, J., Benabbas, L., Bauer, M., Thornhill, N.F. (2009). Cause-and-effect analysis in chemical processes utilizing XML, plant connectivity and quantitative process history. *Computers and Chemical Engineering* , 33, 503-512.
- Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO, *Journal of the Royal Statistical Society B*, 58, 267-288.
- Westerhuis, J.A., Gurden, S.P., Smilde, A.K., (2000). Generalized contribution plots in multivariate statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems*, 51, 95-114.
- Vedam, H., Venkatasubramanian,V. (1999).PCA-SDG based process monitoring and fault diagnosis. *Control Engineering Practice*, 7, 903-917.
- Venkatasubramanian,V., Rengaswamy, R., Yin, K., Kavuri, S.N. (2003a).A review of process fault detection and diagnosis part I: Quantitative model-based methods. *Computers and Chemical Engineering*, 27, 293-311.
- Venkatasubramanian,V., Rengaswamy, R., Kavuri, S.N., Yin, K. (2003b). A review of process fault detection and diagnosis part III: Process history based method, *Computers and Chemical Engineering* , 27, 327-346.

- Yao, Y., Gao, F. (2007). Batch process monitoring in score space of two-dimensional dynamic principal component analysis (PCA), *Industrial and Engineering Chemistry Research*, 46, 8033-8043.
- Yao, Y., Gao, F. (2009). A survey on multistage/multiphase statistical modeling methods for batch processes. *Annual Reviews in Control*, 33, 172-183.
- Yi, G., Shi, J.Q., Choi, T. (2011). Penalized Gaussian process regression and classification for high-dimensional nonlinear data, *Biometrics*, 67, 1285-1294.
- Yoon, S., MacGregor, J.F. (2001). Fault diagnosis with multivariate statistical models part I: using steady state fault signatures, *Journal of Process Control*, 11, 387-400.
- Zadeh, L.A. (1983) The role of fuzzy-logic in the management of uncertainty in expert systems, *Fuzzy Sets and Systems*, 11,199-227.