Journal of Process Control 26 (2015) 17–25

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Canonical variate analysis-based contributions for fault identification

Benben Jiang[a,b], Dexian Huang[a], Xiaoxiang Zhu[b], Fan Yang[a], Richard D. Braatz[b,*]

[a] Department of Automation, Tsinghua University and Tsinghua National Laboratory for Information Science and Technology, Beijing 100084, China
[b] Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

ARTICLE INFO

Article history:
Received 4 August 2014
Received in revised form 18 October 2014
Accepted 3 December 2014
Available online 9 January 2015

Keywords:
Fault identification
Canonical variate analysis
Contribution plot
Tennessee Eastman process
Contribution chart
Process monitoring

ABSTRACT

While canonical variate analysis (CVA) has been used as a dimensionality reduction technique to take into account serial correlations in the process data with system dynamics, its effectiveness in fault identification (i.e., identification of variables most closely associated with a fault) in industrial processes has not been extensively investigated. This paper proposes CVA-based contributions for fault identification, where two types of contributions are developed based on the variations in the canonical state space and in the residual space. The two contributions are used to categorize faulty variables into state-space faulty variables (SSFVs) and residual-space faulty variables (RSFVs), which enhances the understanding of the character of each fault as well as the performance of fault monitoring based on different statistics. The effectiveness of the proposed approach is demonstrated on the Tennessee Eastman process. The simulation results show that the faulty variables identified by the CVA-based contributions can impact the statistics of the state space, the residual space, or both; and abnormal events are observed to be more often linked to faulty variables in the residual space rather than in the state space.

1. Introduction

A *fault* is defined as any abnormal event that occurs during process operations. Investigating the causes of faults is critical for the efficient and optimal operation of industrial processes. As manufacturing facilities become increasingly integrated and large-scale – largely due to efforts to reduce energy costs, maximize profit, and reduce environmental releases – the potential for faults to dynamically propagate in nonintuitive ways to produce significant harm to equipment, life, and the environment has increased. These trends motivate the development of methods to quickly identify variables associated with a fault as quickly as possible, preferably before its effects propagate to the extent of becoming a major safety concern. Such methods are collectively referred to as *fault identification*, in contrast to *fault detection*, which is the first step in a data-based process monitoring scheme that detects whether some fault has occurred, and *fault diagnosis*, in which the specific cause of the fault is determined.

Accurate first-principles dynamic models do not exist for most manufacturing facilities, which is why existing process monitoring systems in industry are typically constructed based on measured data collected and stored in a historical database. In this approach, information about the process operations needs to be abstracted from the historical data. This task of fault identification can be rather challenging when there are a large number of strongly correlated process variables, as is typical in most chemicals, petrochemicals, and refining operations.

The proficiency of identifying faults from data can be improved using dimensionality reduction techniques, such as principal component analysis (PCA) [1,2], partial least-squares (PLS) analysis [3,4], and canonical variate analysis (CVA) [5–7]. In statistical process monitoring, the PCA and PLS methods has been observed to perform well for process measurements when the in-control variations are independent and identically distributed (i.i.d.). With the i.i.d. assumption, the entire variability of the observations can be explained by estimating the covariance without time lags. If the vectors of observations are serially correlated, that is, the observations at one time instant are correlated with those at past time instants, the zero-lag covariance matrix cannot fully represent the entire variation.

To handle serially correlated multivariate observations, PCA [8] and PLS techniques [9] have been utilized by constructing a covariance matrix with time lags. A dynamic model is extracted directly from data by performing the time lagged version of PCA [10,11]. Utilizing the eigenvector of the covariance matrix that corresponds to the zero eigenvalue, a multivariate autoregressive with exogenous input (ARX) model is developed in [10]. The eigenvector corresponding to a nearly zero eigenvalue is an approximate representation of the ARX model [5,10]. A drawback of this approach

* Corresponding author at: 77 Massachusetts Avenue, Room E19-551, Cambridge, MA 02139, USA. Tel.: +1 617 253 3112; fax: +1 617 258 0546.
E-mail address: braatz@mit.edu (R.D. Braatz).

http://dx.doi.org/10.1016/j.jprocont.2014.12.001
0959-1524/

is the inflexibility of the ARX model for the description of dynamic processes [12].

Aside from methods derived from PCA and PLS, CVA is also a dimensionality reduction technique that selects pairs of variables from the inputs and outputs that maximizes a correlation statistic [12–14]. This method takes serial correlations into account during the dimensionality reduction procedure. In the CVA approach, the statistical model extracted by the CVA approach is in the form of a state-space model where the state variables are statistically independent at zero lag [5].

Contribution plots [15] are the most popular technique for determining which variables are most strongly associated with the statistics no longer being within the normal operating condition (NOC). A higher contribution of a process variable indicates that the fault-related deviations in the specific variable are larger. Several efforts have been published that identified weaknesses and/or proposed modifications of contribution charts to improve their ability to identify the variables of most value in identifying the location of the fault (aka "faulty variables"). Utilizing a scheme of PCA-based contribution plot, Kourti and MacGregor [16] determined faulty variables of a high-pressure low-density polyethylene reactor, and reported that the PCA-based contribution plots may not always correctly identify the most important variables associated with a fault. Yue and Qin [17] proposed an index that combines T^2 and Q statistics during fault identification, which was shown to be more effective than using a single statistic [18]. By introducing confidence limits into contribution plots, Westerhuis et al. [19] improved the statistical analysis of faulty variables, and concluded that the contribution plots should be carefully interpreted due to a smearing effect in the residuals of PCA. Since the smearing effect may mislead the determination of faulty variables, Liu [20] proposed contribution plots without the smearing effect on non-faulty variables by maximizing the reduction of a combined index through a missing data method. The data-driven and model-based methods for fault identification were comprehensively compared in [21], which reported that the identification of simple faults can be easily provided by contribution plots, while the determination of complicated faults needs additional information about the process operations. Furthermore, a reconstruction-based approach for determining faulty variables from the subspace of abnormal events is developed in [18,22], where an identification index is utilized to identify faults. The identification index is defined as the ratio of the reconstructed squared prediction error (SPE) to the faulty SPE.

Both PCA- and PLS-based contribution plots are limited in their ability to quickly identify faults because the underlying PCA and PLS methods do not produce the most accurate dynamic models, even when lagged data are used in their construction. This drawback has been recognized, and contribution plots in conjunction with state-space models have been carried out to better take into account the process dynamics [23–26,29]. In a few studies [23–25], subspace system identification based on N4SID has been utilized to obtain a state-space model that was used to construct contribution plots. Although CVA has been demonstrated to outperform N4SID for subspace identification in terms of stability and parsimony (fewer parameters) in the representation of dynamic systems [26–28], investigation on the application of contribution plots to CVA is limited. In one study where contribution plots and CVA-based state-space models were investigated [29], contributions were calculated based on the statistics of the states in the state-space model. In the study [29], process inputs were not considered in the state-space model.

In this article, the contribution plots are proposed to be applied to both the state space (retained states in the state-space model obtained via CVA) and the residual space (the rest of the states in the CVA model), to examine the utility of the two different measures for fault identification. The two types of contributions (state space and residual space) correspond to different characteristics of the process and can potentially provide more insights into the fault. This article plots the contributions as two-dimensional color maps, which allows improved fault identification compared to the traditional one-dimensional plots [30].

The rest of this paper is organized as follows. The CVA approach is briefly described in Section 2. The contribution maps for CVA-based state and residual spaces are developed in Section 3. The effectiveness of the proposed scheme is demonstrated in the Tennessee Eastman process in Section 4, followed by conclusions in Section 5.

2. Canonical variate analysis (CVA) revisited

2.1. The CVA statistical method

CVA is a dimensionality reduction technique in multivariate statistical analysis, which maximizes a correlation statistic with selected two sets of variables. Assuming process input and output vectors $\boldsymbol{x} \in R^m$ and $\boldsymbol{y} \in R^n$ with covariance matrices $\boldsymbol{\Sigma_{xx}}$ and $\boldsymbol{\Sigma_{yy}}$ and cross-covariance matrix $\boldsymbol{\Sigma_{xy}}$, matrices $\boldsymbol{J} \in R^{m\times m}$ and $\boldsymbol{L} \in R^{n\times n}$ can be obtained under the condition that

$$\begin{cases} \boldsymbol{J}\boldsymbol{\Sigma}_{xx}\boldsymbol{J}^{\mathrm{T}} = \boldsymbol{I}_{\overline{m}} \\ \boldsymbol{L}\boldsymbol{\Sigma}_{yy}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{I}_{\overline{n}} \end{cases} \tag{1}$$

and

$$\boldsymbol{J}\boldsymbol{\Sigma}_{xy}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{D} = \mathrm{diag}(\gamma_1, \ldots, \gamma_r, 0, \ldots, 0) \tag{2}$$

where $\overline{m} = \mathrm{rank}(\boldsymbol{\Sigma_{xx}})$, $\overline{n} = \mathrm{rank}(\boldsymbol{\Sigma_{yy}})$, $r = \mathrm{rank}(\boldsymbol{\Sigma_{xy}})$, γ_i ($i = 1, 2, \ldots, r$) are canonical correlations with $\gamma_1 \geq \cdots \geq \gamma_r$, and $\boldsymbol{I}_k$ is a block-diagonal matrix with a $k \times k$ identity matrix as the first block and a zero matrix as the second block [13]. The vectors of canonical variables $\boldsymbol{c} = \boldsymbol{J}\boldsymbol{x}$ and $\boldsymbol{d} = \boldsymbol{L}\boldsymbol{y}$ contain a set of independent variables with the covariance matrix $\boldsymbol{\Sigma_{cc}} = \boldsymbol{I}_{\overline{m}}$ and $\boldsymbol{\Sigma_{dd}} = \boldsymbol{I}_{\overline{n}}$, respectively.

By solving the singular value decomposition (SVD)

$$\sum_{\boldsymbol{xx}}^{-1/2} \sum_{\boldsymbol{xy}} \sum_{\boldsymbol{yy}}^{-1/2} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}, \tag{3}$$

the matrix of canonical correlations $\boldsymbol{D}$ can be computed as $\boldsymbol{D} = \boldsymbol{\Sigma}$; and the projection matrices $\boldsymbol{J}$ and $\boldsymbol{L}$ can be obtained as $\boldsymbol{J} = \boldsymbol{U}^T \boldsymbol{\Sigma}_{\boldsymbol{xx}}^{-1/2}$ and $\boldsymbol{L} = \boldsymbol{V}^T \boldsymbol{\Sigma}_{\boldsymbol{yy}}^{-1/2}$, respectively. The matrices $\boldsymbol{U}^{\mathrm{T}}$ and $\boldsymbol{V}^{\mathrm{T}}$ rotate the canonical variables to be pairwise correlated, and the matrices $\boldsymbol{\Sigma}_{\boldsymbol{xx}}^{-1/2}$ and $\boldsymbol{\Sigma}_{\boldsymbol{yy}}^{-1/2}$ scale the canonical variables to be unit variance.

2.2. CVA state vector

Hotelling proposed the CVA concept for multivariate statistical analysis, but CVA was not utilized for stochastic realization theory and system identification until Akaike's work on ARMA models [13]. The CVA approach was further developed using state-space models by Larimore [13]. Given time series output data $\boldsymbol{y}(t) \in R^{m_y}$ and input data $\boldsymbol{u}(t) \in R^{m_u}$, the linear state-space model is given by [31,33]

$$\boldsymbol{x}(t+1) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t) + \boldsymbol{v}(t) \tag{4}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}\boldsymbol{x}(t) + \boldsymbol{D}\boldsymbol{u}(t) + \boldsymbol{E}\boldsymbol{v}(t) + \boldsymbol{w}(t) \tag{5}$$

where $\boldsymbol{x}(t) \in R^d$ is a d-dimensional state vector, $\boldsymbol{v}(t)$ and $\boldsymbol{w}(t)$ are independent white noise processes, and $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, and $\boldsymbol{E}$ are coefficient matrices.

At a particular time instant $t = 1, 2, \ldots$, the vector including the past information is given by

$$\boldsymbol{p}(t) = [\boldsymbol{y}^{\mathrm{T}}(t-1), \boldsymbol{y}^{\mathrm{T}}(t-2), \ldots, \boldsymbol{u}^{\mathrm{T}}(t-1), \boldsymbol{u}^{\mathrm{T}}(t-2), \ldots]^{\mathrm{T}} \tag{6}$$

and the vector including the present and future output information is given by

$$\boldsymbol{f}(t) = [\boldsymbol{y}^{\mathrm{T}}(t), \boldsymbol{y}^{\mathrm{T}}(t+1), \ldots]^{\mathrm{T}}. \tag{7}$$

By substituting the matrix $\boldsymbol{\Sigma}_{\boldsymbol{xy}}$ with $\boldsymbol{\Sigma}_{\boldsymbol{pf}}$, $\boldsymbol{\Sigma}_{\boldsymbol{xx}}$ with $\boldsymbol{\Sigma}_{\boldsymbol{pp}}$, and $\boldsymbol{\Sigma}_{\boldsymbol{yy}}$ with $\boldsymbol{\Sigma}_{\boldsymbol{ff}}$, the matrices $\boldsymbol{J}$, $\boldsymbol{L}$, and $\boldsymbol{D}$ can be computed via the SVD as Eq. (3). Under the condition that the order d is selected to be equal to or greater than the order of the minimal state-space realization of the actual process, the states $\boldsymbol{x}(t)$ in Eqs. (4) and (5) can be derived from the data using the CVA states as

$$\boldsymbol{x}_d(t) = \boldsymbol{J}_d\boldsymbol{p}(t) = \boldsymbol{U}_d^T\widehat{\boldsymbol{\Sigma}}_{pp}^{-1/2}\boldsymbol{p}(t) \tag{8}$$

where $\boldsymbol{J}_d = \boldsymbol{U}_d^T\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{pp}}^{-1/2}$ and $\boldsymbol{U}_d$ contains the first d columns of $\boldsymbol{U}$ in Eq. (3) [31].

Due to the finite quantity of data available in practice, the vectors $\boldsymbol{p}(t)$ and $\boldsymbol{f}(t)$ are usually truncated to be

$$\boldsymbol{p}(t) = [\boldsymbol{y}^{\mathrm{T}}(t-1), \boldsymbol{y}^{\mathrm{T}}(t-2), \ldots, \boldsymbol{y}^{\mathrm{T}}(t-l), \boldsymbol{u}^{\mathrm{T}}(t-1), \boldsymbol{u}^{\mathrm{T}}(t-2), \ldots, \boldsymbol{u}^{\mathrm{T}}(t-l)]^T \tag{9}$$

and

$$\boldsymbol{f}(t) = [\boldsymbol{y}^{\mathrm{T}}(t), \boldsymbol{y}^{\mathrm{T}}(t+1), \ldots, \boldsymbol{y}^{\mathrm{T}}(t+h)]^{\mathrm{T}} \tag{10}$$

where l and h are the number of lags in vectors $\boldsymbol{p}(t)$ and $\boldsymbol{f}(t)$, respectively. Moreover, the state order of the system is not known in advance. One solution is to fit the data to autoregressive exogenous (ARX) models with several different numbers of lags by least squares and select l and h as the lags that minimize the Akaike information criterion [14].

2.3. Process monitoring measures

In CVA, two types of statistics are calculated from the process model in normal operation [31,32]. The first statistic, which measures the variations inside the canonical state space, is defined as

$$T_d^2(t) = \boldsymbol{x}_d^{\mathrm{T}}(t)\boldsymbol{x}_d(t). \tag{11}$$

The second statistic, which measures the variations in the residual space, is defined as

$$T_e^2(t) = \boldsymbol{x}_e^{\mathrm{T}}(t)\boldsymbol{x}_e(t) \tag{12}$$

where

$$\boldsymbol{x}_e(t) = \boldsymbol{J}_e\boldsymbol{p}(t) = \boldsymbol{U}_e^T\widehat{\boldsymbol{\Sigma}}_{pp}^{-1/2}\boldsymbol{p}(t) \tag{13}$$

and $\boldsymbol{U}_e$ contains the last $e = l(m_u + m_y) - d$ columns of $\boldsymbol{U}$ in Eq. (3). CVA-based statistics for fault detection are also reported in [32,33].

A violation of the T_d^2 statistic (e.g., exceeding the threshold corresponding to the 99% confidence level [31]) indicates that the states are out of control (that is, are in an abnormal operating condition), whereas a violation of the T_e^2 statistic indicates that the characteristic of the measurement noise has changed and/or new states have been created in the process.

3. CVA-based contributions

In this section, the T_d^2 and T_e^2 CVA statistics in Section 2 are used for computing contributions. Large deviations in the variables from normal operating conditions result in large values of the two statistics. While the states $\boldsymbol{x}_d(t)$ alone do not provide a direct way to identify the variables that contribute to large statistics, the loadings ($\boldsymbol{J}_d$ or $\boldsymbol{J}_e$) indicate how the variables contribute to the states in the model and can be used to compute *variable contributions*. First consider the variable contribution based on the state space for the CVA model (the residual space is considered later).

For a new time instant data vector $\boldsymbol{p}(t)$, the contribution based on the state space in the CVA model is defined as

$$\begin{aligned} c^d(t) &= \boldsymbol{x}_d^{\mathrm{T}}(t)\boldsymbol{x}_d(t) = \boldsymbol{x}_d^{\mathrm{T}}(t)(\boldsymbol{J}_d\boldsymbol{p}(t)) \\ &= \boldsymbol{x}_d^T(t)\sum_{k=1}^{(m_y+m_u)l}(p_k(t)\boldsymbol{J}_{d,k}^{\mathrm{T}})^{\mathrm{T}} \\ &= \sum_{k=1}^{(m_y+m_u)l}\boldsymbol{x}_d^{\mathrm{T}}(t)(p_k(t)\boldsymbol{J}_{d,k}^{\mathrm{T}})^{\mathrm{T}} \\ &= \sum_{k=1}^{(m_y+m_u)l}c_{p_k}^d \end{aligned} \tag{14}$$

where $p_k(t)$ is the jth element of the data vector $\boldsymbol{p}(t)$, $\boldsymbol{J}_{d,k}^T$ is the kth row of the $\boldsymbol{J}_d^T$, and

$$c_{p_k}^d(t) = \boldsymbol{x}_d^{\mathrm{T}}(t)(p_k(t)\boldsymbol{J}_{d,k}^{\mathrm{T}})^{\mathrm{T}} \tag{15}$$

is the contribution of $p_k(t)$ in the new time instant vector $\boldsymbol{p}(t)$.

A major difference from PCA is that the vector $\boldsymbol{p}(t)$ in CVA contains serial correlation, that is, the observations at one time instant are correlated with the observations at past time instants.

From (9), $p_k(t)$ can be specifically written as

$$p_k(t) = \begin{cases} y_i(t-j), \text{ and } k = (i-1)m_y + j & \text{when } k \le m_y l \\ u_i(t-j), \text{ and } k = m_y l + (i-1)m_u + j & \text{when } k > m_y l \end{cases} \tag{16}$$

Therefore the contribution based on the state space for variable y_m is the multiplication of all its l lagged observations $y_m(t-j)$, ($j = 1, 2, \ldots, l$):

$$c_{y_m}^d(t) = \sum_{j=1}^{l}\boldsymbol{x}_d^{\mathrm{T}}(t)(y_m(t-j)\boldsymbol{J}_{d,m_j}^{\mathrm{T}})^{\mathrm{T}} \tag{17}$$

The contribution of the process variable u_m is calculated in the same way as for y_m.

Similarly, the contribution of the process variable y_m based on the residual space can be obtained as

$$c_{y_m}^e(t) = \sum_{j=1}^{l}\boldsymbol{x}_e^{\mathrm{T}}(t)(y_m(t-j)\boldsymbol{J}_{e,m_j}^{\mathrm{T}})^{\mathrm{T}}. \tag{18}$$

The contribution of the process variable u_m based on the residual space is also calculated in the same way as for y_m.

A limitation of the CVA-based contributions is that contributions can be overly sensitive due to the inversion of $\widehat{\Sigma}_{pp}^{-1/2}$ in Eqs. (8) and (13) [31], which can result in identifying incorrect faulty variables. This sensitivity can be reduced by combining the two CVA-based contributions (state space c^d and residual space c^e) for fault identification, in a similar manner as has been applied to PCA and PLS [17,18].

A higher contribution of a process variable indicates a more severe abnormal value with the specific variable. Faulty variables identified by contribution plot based on state space (c^d) are called *state-space faulty variables* (SSFVs), i.e., are associated with large deviations of the states that are present during normal operating conditions. The models for those faulty variables are still valid. Faulty variables identified by contribution plot based on residual space (c^e) are called *residual-space faulty variables* (RSFVs). Such variables are associated with new states created in the process, which are no longer sufficiently described by the CVA model constructed for normal operating conditions. Section 4 plots the variable contributions as a two-dimensional color map, which has variable number and sampling instance (aka time) as two axes and

the value of the contribution shown as color. Such 2D contribution maps have been demonstrated [30] to produce more reliable fault identification than the commonly used 1D contribution chart, which has variable number and the value of the contribution as two axes.

4. Application to the Tennessee Eastman process

The Tennessee Eastman process (TEP) is a well-known benchmark process used for comparison of various fault identification methods. The TEP was created to provide a realistic industrial process for control and monitoring studies. The process variables are listed in Table 1 and the TEP flowsheet is shown in Fig. 1. A detailed process description including the specific plantwide closed-loop control system can be found in the textbook [12], which also includes citations to numerous papers on process control and monitoring systems applied to the TEP.

The testing data set for each fault contains 960 observations. The observations were sampled every 3 min. Each data set starts with no faults, and the faults occur after the 160th sample. A total of 52 variables including all the manipulated and measured variables except the agitation speed of the reactor's stirrer were recorded. By using the preprogrammed faults (Faults 1–21), twenty-one testing sets were generated; Fault 0 (with no faults) was generated under normal operating conditions.

In order to avoid particular variables inappropriately dominating the dimensionality reduction procedure, all the data were scaled in the standard manner that is often called "z-scoring". A CVA model was built using a normal operating dataset.

4.1. Fault 1

Fault 1 affects many variables with interacting and counteracting effects that can confuse fault identification methods, and is used to compare the proposed CVA-based contribution plots to the optimized PCA-based contribution plot of [20]. In the Fault 1 scenario, the composition of C in Stream 4 was increased from 51 mol% to 54 mol%; meanwhile, the composition of A of Stream 4 was decreased from 48.5 mol% to 45.5 mol%. These changes lead to a decrease of A in Stream 5, which increases the A feed in Stream 1 by the regulation of a control loop. Subsequently, the variations in the reactor level result from the changes in the compositions and flow rates of Stream 6, and the reactor level's variations affect the flow rate in Stream 4, due to a cascade control loop. The variations of composition C in the reactor feed flow result in the changes of composition E, because of the equimolar reaction among compositions C and E. The variations in the reactants A, C, and E eventually impact the ratio of the products associated with the corresponding material balances.

CVA-based fault detection statistics based on either state space or residual space are very effective at detecting Fault 1 (see Table S1). For fault identification, the CVA-based contributions based on state space (c^d) and residual space (c^e) are displayed in Figs. 2ab as color maps [30]. Both figures show a lot of "noise," which can be explained by the over-sensitivity of CVA-based contributions relying on a single statistic, due to the inversion of $\widehat{\Sigma}_{pp}^{-1/2}$. As mentioned in Section 3, a combined contribution (e.g., $c=(c^d+c^e)/2$ was used in this case study) is recommended and is shown in Fig. 2c for Fault 1.

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jprocont.2014.12.001.

During the early stage (first 400 samples) of Fault 1, numerous faulty variables are reported, as shown in Fig. 2c. With the controllers implemented on the process, most of the faulty variables stabilized to NOC after around the 400th sample (see Fig. 2c). Fig. 3 displays the trends of x_1, x_4, x_{18}, x_{19}, x_{34}, x_{44}, x_{45}, and x_{50}, which are all faulty variables reported by the CVA-based combined contributions (Fig. 2c). The resulting faulty variables are similar with those of [20] except for variable x_{34}, which was not identified by Liu's optimized PCA-based method [20]. The trend of the variable x_{34} is displayed in Fig. 3e. The variable x_{34} is observed to eventually settle to a steady-state value higher than its value at the NOC. The CVA-based contributions correctly identify x_{34} as a faulty variable, which was missed by the optimized PCA-based method [20]. As shown in Fig. 2c, the contributions of the quartet $\{x_1, x_{18}, x_{44}, x_{50}\}$ are much more persistent than the rest of the faulty variables $\{x_4, x_{19}, x_{34}, x_{45}\}$. From Figs. 2ab, the CVA-based contribution plots indicate that x_{18}, x_{19}, and x_{50} are SSFVs, and x_1, x_4, x_{34}, x_{44}, and x_{45} are RSFVs.

The first variable to have a large combined contribution, x_{16}, is the stripper pressure, which directly receives Stream 4 that contains A/B/C feed with the incorrect A/B ratio that defines Fault 1. The four variables with persistent large combined contributions, $\{x_1, x_{18}, x_{44}, x_{50}\}$, are the A feed flow rate in Stream 1, the stripper temperature, the manipulated variable to the A feed flow in Stream 1, and the stripper steam valve. Two of the variables are associated with the stripper, which directly receives the flow from the faulty stream 4, and two of these variables are associated with component A. Collectively these variables would provide some guidance to an operator as to the root cause of the fault.

4.2. Fault 4

The Fault 4 scenario is that a step change in the reactor cooling water inlet temperature results in a step change in the reactor cooling water flow rate (x_{51}), as shown in Fig. 4a. This fault was selected as an example of a fault where the identification of the variables most closely associated with the fault is difficult based on state-space contributions but readily identified based on the residual-space contributions. When the fault occurs at the 160th sample, the reactor temperature (x_9) suddenly increases, as shown in Fig. 4b, which is compensated by the control loops. After some delay, the CVA-based state-space contributions for Fault 4 show large transient deviations in several variables (see the top portion of Fig. 5a), as the effects of the fault propagate to many of the variables associated with the reactor and to some of the variables associated with downstream unit operations. The control systems are able to compensate for the effects of the fault on the process operations, so that the only persistent change in the 51 variables is the reactor cooling water flow rate x_{51}, and the only contribution that persists is x_{51} in the residual space (Fig. 5b).

The CVA-based contributions based on state space (c^d) and residual space (c^e) are displayed in Fig. 5a and b. The combined CVA-based contributions are shown in Fig. 5c, which indicates that variable x_{51} is the faulty variable for Fault 4, which agrees with the time-domain plots (e.g., Fig. 4a and b). The faulty variable x_{51} is an RSFV, as shown in the residual-space contributions in Fig. 5b. Minimal disruption from normal operating conditions is observed in the state space, where Fault 4 results in a large deviation for one variable in the residual space. These results are consistent with results obtained for fault detection based on the CVA model [31, see Table S1]: poor performance (missed detection rate: 0.688) by using the state-space statistic T_d^2 versus excellent performance (missed detection rate: 0) by using the residual-space statistic T_e^2 for the detection of Fault 4.

This example demonstrates the necessity of combined utilization of the two types of CVA-based contributions, whereas utilizing the single state-space contribution can make it impossible to identify faulty variables.

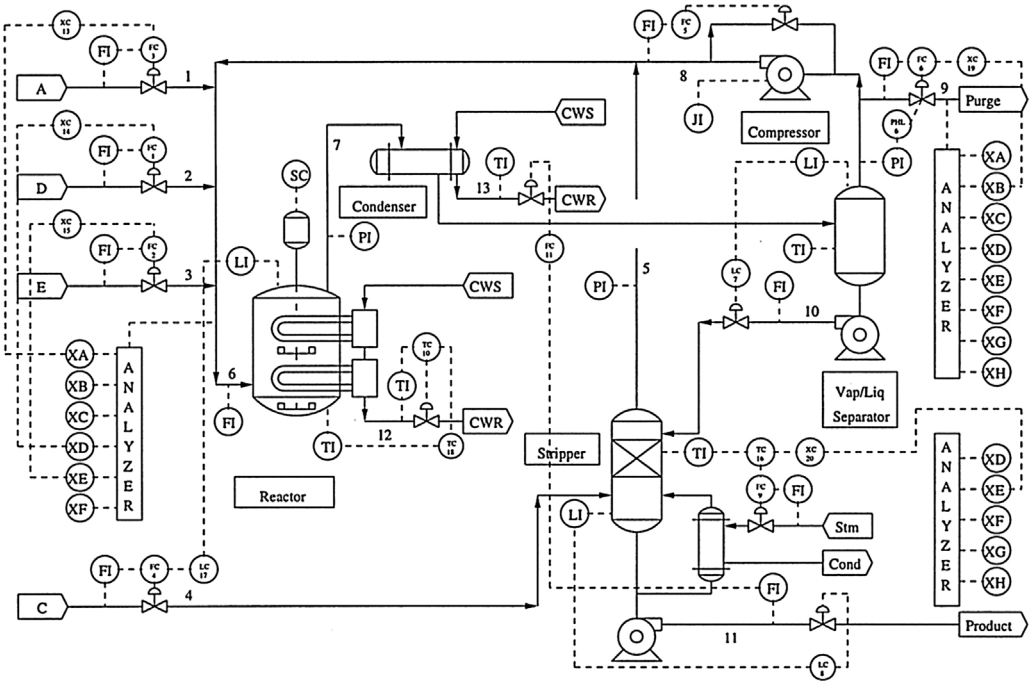

Fig. 1. Flowsheet for the Tennessee Eastman process [31].

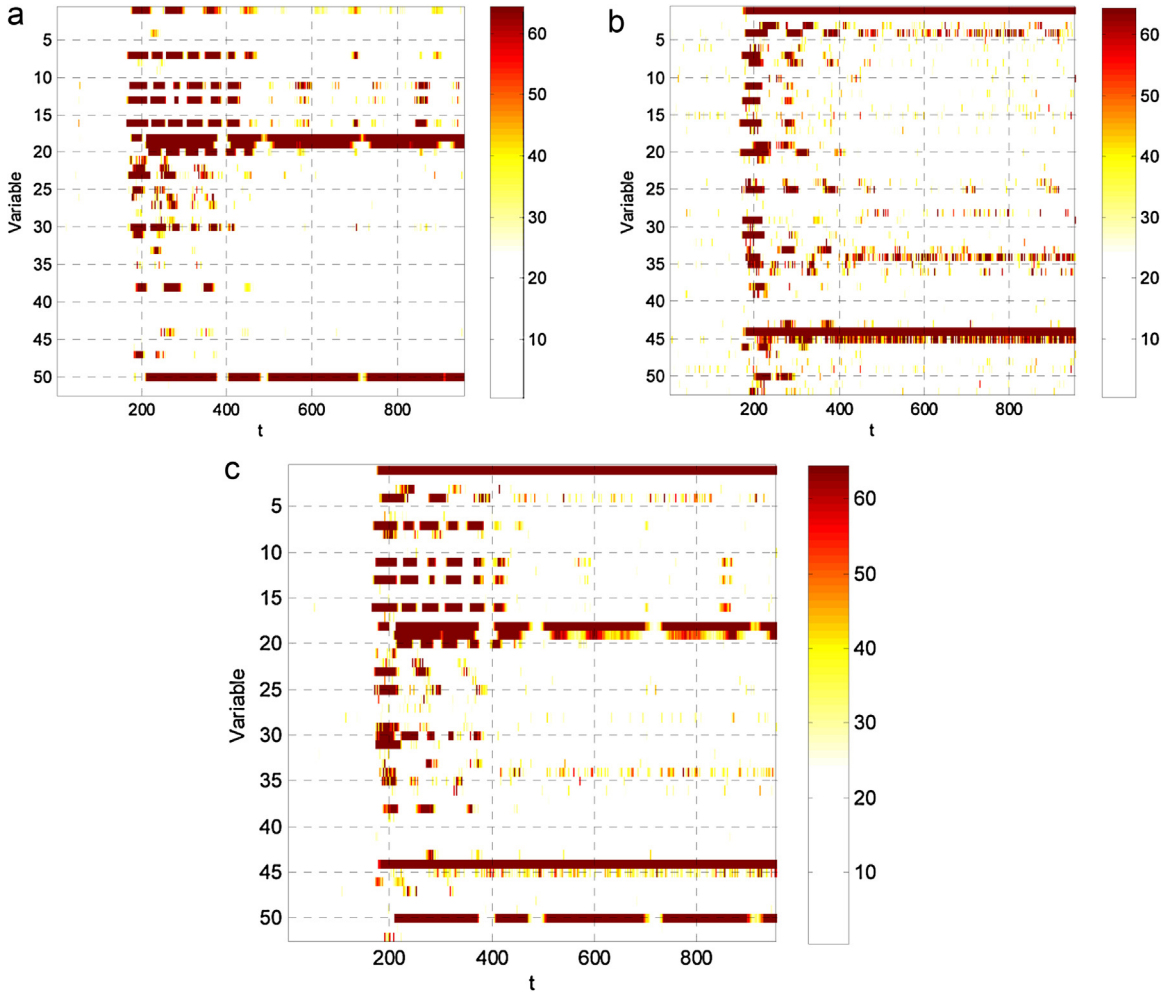

Fig. 2. (a). CVA-based contributions based on state space for Fault 1. (b) CVA-based contributions based on residual space for Fault 1. (c) CVA-based combined contributions for Fault 1. (For interpretation of the references to color in the text, the reader is referred to the web version of the article.)

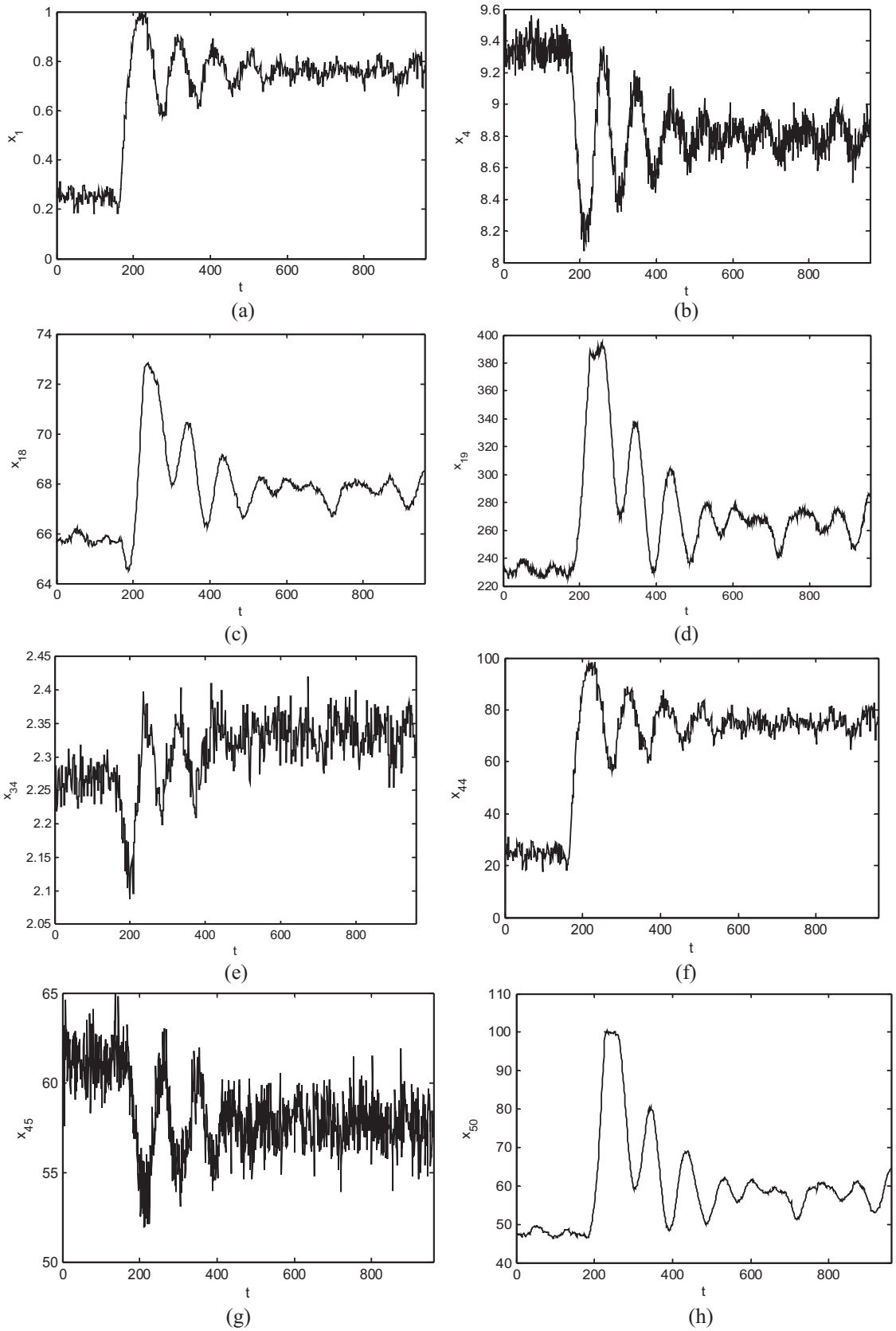

Fig. 3. Trends of the faulty variables for Fault 1 reported by CVA-based contributions. (a) x_1, (b) x_4, (c) x_{18}, (d) x_{19}, (e) x_{34}, (f) x_{44}, (g) x_{45}, and (h) x_{50}.

Table 1
Monitored variables in the Tennessee Eastman process [20].

ID	Variable description	ID	Variable description
x_1	A feed (Stream 1)	x_{27}	Component E (Stream 6)
x_2	D feed (Stream 2)	x_{28}	Component F (Stream 6)
x_3	E feed (Stream 3)	x_{29}	Component A (Stream 9)
x_4	A and C feed (Stream 4)	x_{30}	Component B (Stream 9)
x_5	Recycle flow (Stream 8)	x_{31}	Component C (Stream 9)
x_6	Reactor feed rate (Stream 6)	x_{32}	Component D (Stream 9)
x_7	Reactor pressure	x_{33}	Component E (Stream 9)
x_8	Reactor level	x_{34}	Component F (Stream 9)
x_9	Reactor temperature	x_{35}	Component G (Stream 9)
x_{10}	Purge rate (Stream 9)	x_{36}	Component H (Stream 9)
x_{11}	Product separator temperature	x_{37}	Component D (Stream 11)
x_{12}	Product separator level	x_{38}	Component E (Stream 11)
x_{13}	Product separator pressure	x_{39}	Component F (Stream 11)
x_{14}	Product separator underflow (Stream 10)	x_{40}	Component G (Stream 11)
x_{15}	Stripper level	x_{41}	Component H (Stream 11)
x_{16}	Stripper pressure	x_{42}	MV to D feed flow (Stream 2)
x_{17}	Stripper underflow (Stream 11)	x_{43}	MV to E feed flow (Stream 3)
x_{18}	Stripper temperature	x_{44}	MV to A feed flow (Stream 1)
x_{19}	Stripper stream flow	x_{45}	MV to total feed flow (Stream 4)
x_{20}	Compressor work	x_{46}	Compressor recycle valve
x_{21}	Reactor cooling water outlet temperature	x_{47}	Purge value (Stream 9)
x_{22}	Separator cooling water outlet temperature	x_{48}	Separator pot liquid flow (Stream 10)
x_{23}	Component A (Stream 6)	x_{49}	Stripper liquid product flow (Stream 11)
x_{24}	Component B (Stream 6)	x_{50}	Stripper steam valve
x_{25}	Component C (Stream 6)	x_{51}	Reactor cooling water flow
x_{26}	Component D (Stream 6)	x_{52}	Condenser cooling water flow

4.3. Other faults

The remaining 19 faults were investigated. The orders and the number of lags l determined for CVA are similar to [31]. The CVA-based fault detection results for all faults are reported in Table S1 and the associated fault identification results utilizing CVA-based contribution plot approach are summarized in Table 2, where faulty variables for each fault are categorized into SSFV and RSFV.

Faults 1, 2, 5, 6, 8, 10, 12, 13, 14, 16, 17, 18, and 20 alter the properties of both the state space and residual space, which results in both the contributions c^d and c^e being useful for identifying faulty variables. The attributes of Faults 3, 9, 15, and 21 are not captured by either the state space or the residual space. It has been observed in past studies that these faults are very difficult to detect using any method based solely on historical data (e.g., see Table 4 of [31]). It should be no surprise that identifying faulty variables is nearly impossible for faults that are very difficult to detect, as such faults have very small effects on the measured data.

Faults 4, 7, 11, and 19 can be grouped as a third class of faults, in which only the residual space is significantly affected (see Table 2). For such faults, good fault identification performance is achieved for the residual-space contributions whereas all of the state-space contributions are small.

Except for Faults 10, 12, 16, and 20, the number of faulty variables that occur in the residual space is more than that in the state space. In other words, the abnormal events tend to create faulty variables of RSFV rather than SSFV. The higher chance of faulty variables occurring in the residual space is an indication that the characteristic of the measurement noise has changed and/or new states have been created, and the existing states $x_d(t)$ utilized for the model are no longer sufficient for describing the process dynamics or the input-output relationship.

For some faults, some variables, such as x_1 and x_{18} in Fault 1, appear as either SSFVs or RSFVs, but not as both. Some

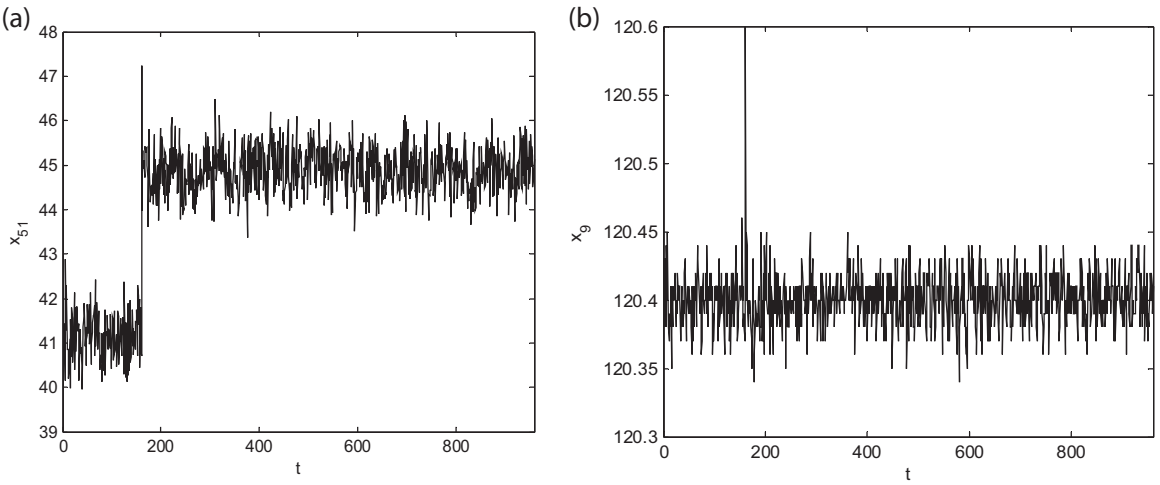

Fig. 4. (a). Trend of the variable x_{51} in Fault 4. (b) Trend of the variable x_9 in Fault 4.

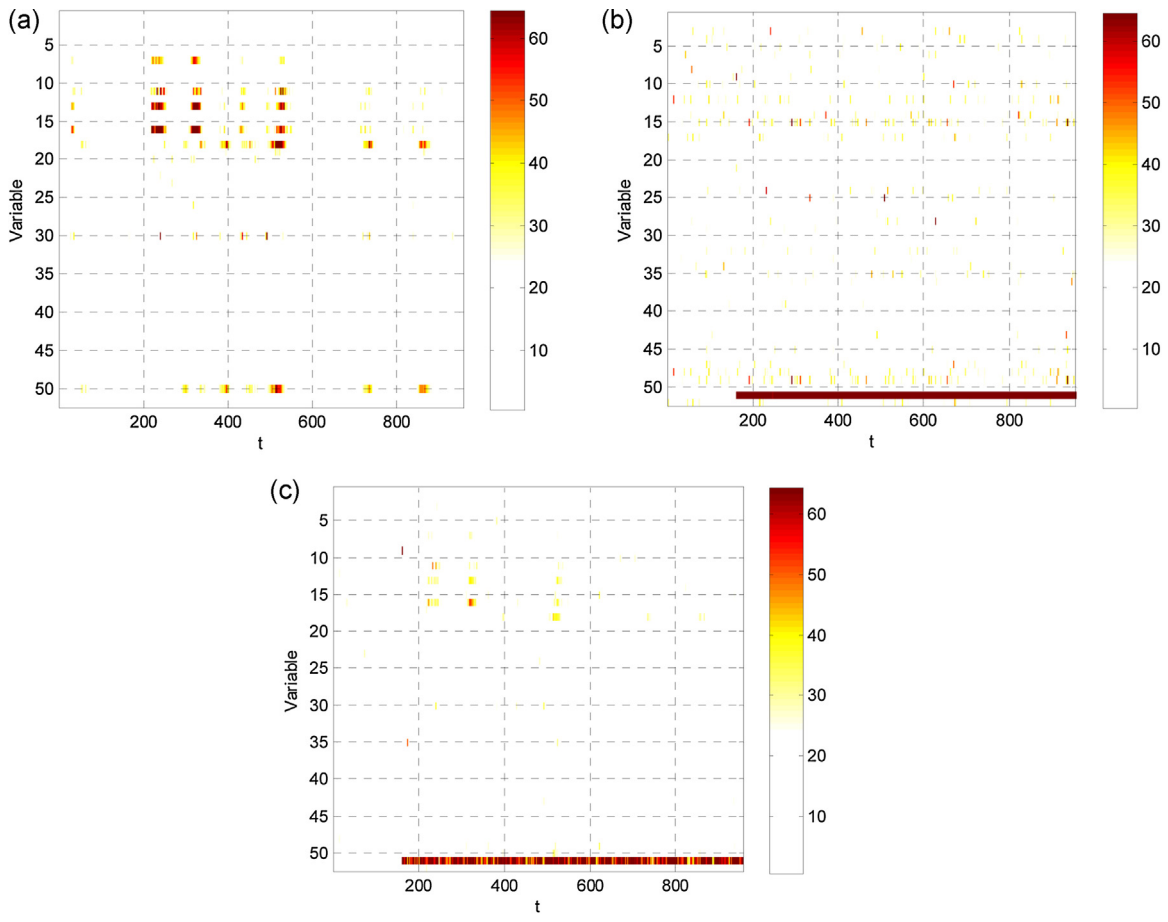

Fig. 5. (a). CVA-based contributions based on state space for Fault 4. (b) CVA-based contributions based on residual space for Fault 4. (c) CVA-based combined contributions for Fault 4.

other variables, such as x_{47} in Fault 2, are both SSFVs and RSFVs, which indicates the faulty variable has significant variations in both the state and the residual spaces of the CVA model.

The relative importance of the state space and residual space on the combined contributions for a particular fault is related to the physical characteristics of the specific fault, in particular, on how the effects of the particular fault propagates through the overall

Table 2
Identification results for all 21 faults.

Fault	SSFV	RSFV
1	x_{18}, x_{19}, x_{50}	$x_1, x_4, x_{34}, x_{44}, x_{45}$
2	$x_{18}, x_{19}, x_{28}, x_{34}, x_{47}, x_{50}$	$x_3, x_9, x_{10}, x_{25}, x_{28}, x_{29}, x_{34}, x_{43}, x_{46}, x_{47}$
3	None	None
4	None	x_{51}
5	x_{50}	x_{52}
6	$x_2, x_3, x_7, x_{11}, x_{13}, x_{16}, x_{21}, x_{22}, x_{23}, x_{25}, x_{26}, x_{30}, x_{31}, x_{33}, x_{35}, x_{36}, x_{38}, x_{43}, x_{44}, x_{47}$	$x_1, x_2, x_3, x_7, x_8, x_9, x_{11}, x_{13}, x_{16}, x_{17}, x_{21}, x_{25}, x_{28}, x_{29}, x_{31}, x_{34}, x_{35}, x_{36}, x_{37}, x_{38}, x_{39}, x_{42}, x_{43}, x_{45}, x_{46}, x_{51}, x_{52}$
7	None	x_{45}
8	$x_7, x_{10}, x_{11}, x_{13}, x_{16}, x_{18}, x_{19}, x_{20}, x_{23}, x_{30}, x_{47}, x_{50}$	$x_1, x_4, x_7, x_{10}, x_{13}, x_{20}, x_{24}, x_{25}, x_{28}, x_{29}, x_{31}, x_{34}, x_{35}, x_{43}, x_{44}, x_{46}, x_{47}$
9	None	None
10	$x_7, x_{13}, x_{19}, x_{50}$	x_{18}
11	None	x_9, x_{51}
12	$x_7, x_9, x_{11}, x_{13}, x_{16}, x_{18}, x_{19}, x_{20}, x_{22}, x_{23}, x_{27}, x_{29}, x_{30}, x_{31}, x_{33}, x_{36}, x_{38}, x_{47}, x_{50}$	$x_1, x_3, x_4, x_8, x_{20}, x_{21}, x_{25}, x_{35}, x_{43}, x_{52}$
13	$x_7, x_{11}, x_{13}, x_{16}, x_{18}, x_{19}, x_{20}, x_{21}, x_{22}, x_{23}, x_{26}, x_{27}, x_{30}, x_{31}, x_{33}, x_{36}, x_{38}, x_{41}, x_{44}, x_{47}, x_{50}$	$x_3, x_4, x_7, x_8, x_{11}, x_{13}, x_{16}, x_{18}, x_{19}, x_{20}, x_{25}, x_{27}, x_{28}, x_{29}, x_{30}, x_{32}, x_{34}, x_{35}, x_{36}, x_{37}, x_{43}, x_{50}, x_{51}$
14	x_{50}	x_9, x_{21}, x_{51}
15	None	None
16	$x_7, x_{13}, x_{18}, x_{19}, x_{20}, x_{50}$	x_{16}
17	$x_7, x_9, x_{11}, x_{13}, x_{16}, x_{18}, x_{19}, x_{21}, x_{30}, x_{50}$	x_9, x_{21}, x_{51}
18	$x_4, x_7, x_9, x_{11}, x_{13}, x_{14}, x_{16}, x_{20}, x_{22}, x_{23}, x_{24}, x_{26}, x_{27}, x_{28}, x_{29}, x_{30}, x_{33}, x_{35}, x_{36}, x_{39}, x_{41}, x_{46}, x_{47}, x_{50}, x_{51}$	$x_6, x_7, x_{12}, x_{13}, x_{16}, x_{17}, x_{19}, x_{21}, x_{23}, x_{25}, x_{27}, x_{28}, x_{30}, x_{32}, x_{33}, x_{35}, x_{36}, x_{37}, x_{38}, x_{41}, x_{42}, x_{43}, x_{45}, x_{46}, x_{48}, x_{51}, x_{52}$
19	None	x_5
20	$x_7, x_{11}, x_{13}, x_{16}, x_{18}, x_{20}, x_{30}, x_{50}$	x_{20}, x_{46}
21	None	None

process system to affect the measured variables and manipulated variables (via the control loops). Some faults have stronger effects on the state space than on the residual space, whereas other faults result in stronger effects on the residual space than the state space. Due the stronger interactions between variables, including the effects of multiple control loops, it is not always easy to assess a priori which fault will have stronger effects on the residual space or the state space. Plausibility arguments can be made based on the process flowsheet as to why the contributions for a particular fault appear more strongly in the state space or residual space, but such plausibility arguments should not be treated as rigorous proofs.

5. Conclusions

This article presents CVA-based contributions for identifying variables that are closely associated with faults, with contributions based on the variations in the state and residual spaces. Faulty variables are categorized into SSFVs and RSFVs, which provide some insight into the character of each fault. The proposed approach has been demonstrated for all the 21 faults in TEP, and the simulation results show that the faulty variables identified by the CVA-based contributions can impact the statistics of the state space, the residual space, or both; and abnormal events are observed to be more often linked to faulty variables in the residual space rather than in the state space.

The combined contributions in the case studies used equal weights for the state and residual spaces ($c = (c^d + c^e)/2$). A consideration for future work is whether the CVA-based fault identification would be improved if different weights were used for the state and residual spaces. Given that variables with large contributions are more likely to be observed in the residual space, should the residual space have a stronger weight, or would this magnify noise? The definition of an optimal weighting may require assuming prior knowledge on the faults such as their probabilities or an upper bound on the number of variables that may be affected by them.

Acknowledgements

This work is supported by the 973 Program of China (Grant No. 2012CB720500) and the National Natural Science Foundation of China (21276137). The first author is grateful for the financial support from China Scholarship Council. The third and fifth authors were supported by the Massachusetts Institute of Technology.

References

[1] A. Raich, A. Cinar, Statistical process monitoring and disturbance diagnosis in multivariable continuous processes, AIChE J. 42 (1996) 995–1009.
[2] J.E. Jackson, G.S. Mudholkar, Control procedures for residuals associated with principal component analysis, Technometrics 21 (1979) 341–349.
[3] J.F. MacGregor, C. Jaeckle, C. Kiparissides, M. Koutoudi, Process monitoring and diagnosis by multiblock PLS methods, AIChE J. 40 (1994) 826–838.
[4] A. Negiz, A. Cinar, On the detection of multiple sensor abnormalities in multivariate processes, in: Proceedings of the American Control Conference, Chicago, 1992, pp. 2364–2368.
[5] A. Negiz, A. Çinar, Statistical monitoring of multivariable dynamic processes with state-space models, AIChE J. 43 (1997) 2002–2020.
[6] P.E.P. Odiowei, Y. Cao, Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations, IEEE Trans. Ind. Inf. 6 (2010) 36–45.
[7] C. Lee, S.W. Choi, I.B. Lee, Variable reconstruction and sensor fault identification using canonical variate analysis, J. Process Control 16 (2006) 747–761.
[8] P. Nomikos, J.F. MacGregor, Monitoring batch processes using multiway principal component analysis, AIChE J. 40 (1994) 1361–1375.
[9] P. Nomikos, J.F. MacGregor, Multi-way partial least squares in monitoring batch processes, Chemom. Intell. Lab. Syst. 30 (1995) 97–108.
[10] W. Ku, R.H. Storer, C. Georgakis, Disturbance detection and isolation by dynamic principal component analysis, Chemom. Intell. Lab. Syst. 30 (1995) 179–196.
[11] B.M. Wise, N.B. Gallagher, The process chemometrics approach to process monitoring and fault detection, J. Process Control 6 (1996) 329–348.
[12] L.H. Chiang, E.L. Russell, R.D. Braatz, Fault Detection and Diagnosis in Industrial Systems, Springer Verlag, London, UK, 2001.
[13] W.E. Larimore, Canonical variate analysis in control and signal processing, in: T. Katayama, S. Sugimoto (Eds.), Statistical Methods in Control and Signal Processing, Marcel Dekker Inc., New York, 1997, pp. 83–120.
[14] W.E. Larimore, Statistical optimality and canonical variate analysis system identification, Signal Processing 52 (1996) 131–144.
[15] P. Miller, R.E. Swanson, Contribution plots: a missing link in multivariate quality control, Appl. Math. Comp. Sci. 8 (1998) 775–792.
[16] T. Kourti, J.F. MacGregor, Multivariate SPC methods for process and product monitoring, J. Qual. Technol. 28 (1996) 409–428.
[17] H.H. Yue, S.J. Qin, Reconstruction-based fault identification using a combined index, Ind. Eng. Chem. Res. 40 (2001) 4403–4414.
[18] R. Dunia, S.J. Qin, Subspace approach to multidimensional fault identification and reconstruction, AIChE J. 44 (1998) 1813–1831.
[19] J.A. Westerhuis, S.P. Gurden, A.K. Smilde, Generalized contribution plots in multivariate statistical process monitoring, Chemom. Intell. Lab. Syst. 51 (2000) 95–114.
[20] J. Liu, Fault diagnosis using contribution plots without smearing effect on non-faulty variables, J. Process Control 22 (2012) 1609–1623.
[21] S. Yoon, J.F. MacGregor, Statistical and causal model-based approaches to fault detection and isolation, AIChE J. 46 (2000) 1813–1824.
[22] R. Dunia, S.J. Qin, Joint diagnosis of process and sensor faults using principal component analysis, Control Eng. Pract. 6 (1998) 457–469.
[23] R.J. Treasure, U. Kruger, J.E. Cooper, Dynamic multivariate statistical process control using subspace identification, J. Process Control 14 (2004) 279–292.
[24] D. Lieftucht, U. Kruger, G.W. Irwin, R.J. Treasure, Fault reconstruction in linear dynamic systems using multivariate statistics, IEE Proc. – Control Theory Appl. 153 (2006) 437–446.
[25] P.K. Li, R.J. Treasure, U. Kruger, Dynamic principal component analysis using subspace model identification, in: D.S. Huang, X.P. Zhang, G.B. Huang (Eds.), Advances in Intelligent Computing, Springer-Verlag, Berlin and Heidelberg, Germany, 2005, pp. 727–736.
[26] Y. Yao, F. Gao, Subspace identification for two-dimensional dynamic batch process statistical monitoring, Chem. Eng. Sci. 63 (2008) 3411–3418.
[27] B.C. Juricek, W.E. Larimore, D.E. Seborg, Reduced rank ARX and subspace system identification for process control, in: Proceedings of the 5th IFAC Symposium on Dynamics and Control of Process Systems, 1998, pp. 247–252.
[28] A. Simoglou, E.B. Martin, A.J. Morris, Statistical performance monitoring of dynamic multivariate processes using state space modeling, Comput. Chem. Eng. 26 (2002) 909–920.
[29] A. Norvilas, A. Negiz, J. DeCicco, A. Cinar, Intelligent process monitoring by interfacing knowledge-based systems and multivariate statistical monitoring, J. Process Control 10 (2000) 341–350.
[30] X. Zhu, R.D. Braatz, Two-dimensional contribution map for fault identification, IEEE Control Syst. 34 (2014) 72–77.
[31] E.L. Russell, L.H. Chiang, R.D. Braatz, Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis, Chemom. Intell. Lab. Syst. 51 (2000) 81–93.
[32] B.C. Juricek, D.E. Seborg, W.E. Larimore, Fault detection using canonical variate analysis, Ind. Eng. Chem. Res. 43 (2004) 458–474.
[33] H.J. Borsje, Fault detection in boilers using canonical variate analysis, in: Proceedings of the American Control Conference, San Diego, CA, 1999, pp. 1167–1170.