Multi-layer Contribution Propagation Analysis for Fault Diagnosis

Ruo-Mu Tan[1]   Yi Cao[2]

[1] School of Water, Energy and Environment, Cranfield University, Cranfield MK43 0AL, UK

[2] College of Chemical and Biological Engineering, Zhejiang University, Hangzhou 310027, China

Abstract: The recent development of feature extraction algorithms with multiple layers in machine learning and pattern recognition has inspired many applications in multivariate statistical process monitoring. In this work, two existing multi-layer linear approaches in fault detection are reviewed and a new one with extra layer is proposed in analogy. To provide a general framework for fault diagnosis in succession, this work also proposes the contribution propagation analysis which extends the original definition of contribution of variables in multivariate statistical process monitoring. In fault diagnosis stage, the proposed contribution propagation analysis for multilayer linear feature extraction algorithms is compared with the fault diagnosis results of original contribution plots associated with single layer feature extraction approach. Plots of variable contributions obtained by the aforementioned approaches on the data sets collected from a simulated benchmark case study (Tennessee Eastman process) as well as an industrial scale multiphase flow facility are presented as a demonstration of the usage and performance of the contribution propagation analysis on multi-layer linear algorithms.

Keywords: Process monitoring, fault detection and diagnosis, contribution plots, feature extraction, multivariate statistics.

1 Introduction

The fast-developing feature extraction approaches in machine learning and pattern recognition, such as neural network and Bayesian network, are playing a more and more important role in data-driven fault detection and diagnosis[1, 2], many of which have multiple layers of data analysis so that the algorithm performance can be enhanced. In the meantime, multivariate statistical process monitoring approaches, such as principle component analysis[3], canonical variate analysis[4], Fisher discriminant analysis[5], independent component analysis[6] and their kin[7–11], may also take the advantage of multiple layers of learning in order to handle the complexities in process data analytics and data-driven process monitoring. Existing complexities, such as nonlinearity, temporal correlation, multimodality, and non-Gaussianity, may be reflected by process data especially when abundant measurements are available for process data analytics while each basic multivariate data analytic method may cope with only some of them individually. Therefore, multi-layer feature extraction structure is of interest for multivariate statistical process monitoring and attention has been paid to this establishment recently. For instance, Cao and Samuel[12] have established a principal component analys-

is PCA-enhanced canonical variate analysis (CVA) method for fault detection and applied it to the Tennessee Eastman challenge process. The state-space independent component analysis (ICA)[13] was developed by combining CVA and ICA in order to handle both temporal correlation and non-Gaussianity in process data. More examples of this establishment include CVA-fisher discriminant analysis (FDA)[9] and ICA-PCA[14] based process monitoring. Similar multi-layer approach has been proposed for nonlinear process monitoring[15, 16] and heterogeneous data fusion[16]. These findings inspire the investigation of improving process monitoring algorithms with extra layers of feature extraction, and the PCA-enhanced state-space ICA approach proposed in this work is an example.

An existing issue with these multi-layer algorithms is their interpretability. The features obtained may be more representative and sensitive to fault occurrence; the transparency of algorithm, however, will be sacrificed due to multiple layers of projection. As an example of data-driven fault diagnosis methods, contribution plots[17] have been extensively applied to identifying variables associating to a certain fault and locating the fault. The contribution plots root in the idea that the process variable that has a significant impact on the fault is supposed to have larger contributions to the monitoring statistics. Therefore, studies and applications of this idea are persistent and profound[18–20]. The mathematical formulation of contribution also provides flexibility for extensions in different scenarios, such as generalized contribution plots[21], reconstruction-based contributions[22] and contri-

bution rate[23]. Nevertheless, how to extract contributions from multi-layer algorithms is still not well understood. This work exploits the usage of contribution plots in faulty variable identification when multi-layer linear feature extraction algorithms are adopted in multivariate statistical process monitoring. A general structure of contribution propagation analysis is proposed accordingly. This analysis takes a step back to the original definition of contribution and investigates the way it propagates between layers with respect to the mathematical formulation of different multi-layer linear feature extraction algorithms so as to quantify the influence of original variables on the final monitoring statistics. In the fault diagnosis stage, once the contributions propagating through the intermediate layers have been quantitatively analyzed, it is reasonable to infer that contribution of influential variables will be properly emphasized and these variables can be identified. Applying the proposed analysis, the contribution propagation solution is derived for the following linear multi-layer feature extraction algorithms: PCA-enhanced CVA, state-space ICA, and PCA-enhanced state-space ICA.

As an extension to a preliminary study on this topic[24], in which a two-layer contribution analysis for PCA-enhanced CVA was presented, this work has the following contributions. Firstly, a novel three-layer feature extraction algorithm, PCA-enhanced state-space ICA, is proposed as a more sophisticated and advanced feature extraction approach; its application in multivariate statistical process monitoring is presented along with PCA-enhanced CVA and state-space ICA. Secondly, the contribution propagation framework is proposed for faulty variable identification by extending the contribution analysis adopted for PCA-enhanced CVA based process monitoring. By applying this framework, the applicability of contribution-based fault diagnosis methods will be enhanced and the interpretability of various multi-layer feature extraction algorithms in process monitoring will be improved significantly. Improved interpretability and faulty variable identification results of proposed PCA-enhanced state-space ICA and contribution propagation analysis can be demonstrated via well-acknowledged benchmark data set from Tennessee Eastman (TE) process; furthermore, these approaches are also feasible for real-life industrial plant monitoring.

The organization of this paper is as follows. The general structure of fault detection algorithms based on multi-layer linear feature extraction is revisited in Section 2.1 along with the formulation of PCA-enhanced CVA, state-space ICA, and PCA-enhanced state-space ICA. For fault diagnosis, contribution propagation analysis is proposed for contribution plots calculation in these multi-layer linear feature extraction based monitoring techniques and the solutions to these contribution plots in Section 3. Section 4 presents two case studies: 1) The simulated case study of Tennessee Eastman process; 2) A case study on the benchmark data set collected from the multiphase

flow facility. The performance of identifying influential variables based on their contributions using the following linear algorithms are compared in succession: CVA with ordinary contribution plots (single-layer)[18]; PCA-enhanced CVA, state-space ICA with contribution propagation (2-layer); and PCA-enhanced state-space ICA with contribution propagation (3-layer). Section 5 summarizes the findings in this work and illustrates potential directions of extension in future study.

2 Multiple layer linear algorithms for process monitoring

Fig. 1 represents the general formulation of multi-layer linear feature extraction based fault detection algorithm.
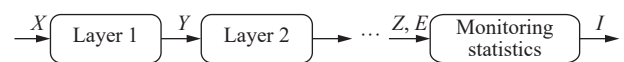

Fig. 1   Algorithm structure

X is the original data, Y_i are the intermediate features obtained by each feature extraction layer, Z and E are the representative features and residuals, respectively. The monitoring statistic I is calculated using Z and E.

PCA-enhanced CVA, state-space ICA and the new PCA-enhanced state-space ICA will be introduced next as three examples of multi-layer linear algorithms shown in Fig. 1.

2.1 PCA-enhanced CVA

According to [12], CVA fault detection algorithm is enhanced by using latent variables extracted by PCA as its input instead of the original measured variables. Firstly, PCA projects the original measurement data in v-dimensional variable space to a reduced r-dimensional principal component space with maximum explanation of variations in original variables. The model structure of PCA is illustrated as follows:

$$Y = XP \qquad (1)$$

where X with zero mean and unit variance is the standardized original data set in the Euclidean space $\mathbf{R}^{n \times v}$, $Y \in \mathbf{R}^{n \times r}$ is the extracted principal components and $P \in \mathbf{R}^{v \times r}$ is the projection matrix. The projection matrix P is obtained by eigenvalue decomposition of sample covariance matrix $X^{\mathrm{T}}X$. Hence, the principal component vector $\boldsymbol{y}$ is linear projection of original variable vector $\boldsymbol{x}$.

CVA is a linear dynamic feature extraction method from which the canonical variates with maximum correlation between past and future vectors can be acquired. Instead of using original data matrix X, PCA-enhanced CVA algorithm adopts the principal components Y extracted by PCA as the input to CVA and get canonical

variate matrix $Z \in \mathbf{R}^{n \times d}$ and residual matrix $E \in \mathbf{R}^{n \times r}$.

At certain time stamp t, past and future vectors $\boldsymbol{y}_p(t)$ and $\boldsymbol{y}_f(t)$ are formed by (2) with fixed vector lengths p and f:

$$\boldsymbol{y}_p(t) = [\tilde{\boldsymbol{y}}^{\mathrm{T}}(t-1), \tilde{\boldsymbol{y}}^{\mathrm{T}}(t-2), \cdots, \tilde{\boldsymbol{y}}^{\mathrm{T}}(t-p)]^{\mathrm{T}}$$
$$\boldsymbol{y}_f(t) = [\tilde{\boldsymbol{y}}^{\mathrm{T}}(t), \tilde{\boldsymbol{y}}^{\mathrm{T}}(t+1), \cdots, \tilde{\boldsymbol{y}}^{\mathrm{T}}(t+f)]^{\mathrm{T}} \qquad (2)$$

where $\tilde{\boldsymbol{y}}(t) = \boldsymbol{y}(t) - \bar{\boldsymbol{y}}$ such that $\bar{\boldsymbol{y}}$ is the mean of principal components $\boldsymbol{y}$ over time.

Furthermore, past and future Henkel matrices $Y_p \in \mathbf{R}^{rp \times m}$ and $Y_f \in \mathbf{R}^{rf \times m}$ comprise of $m = n - p - f + 1$ past/future vector pairs, making the time lagged data matrices for feature extraction starting at time stamp p (the minimal initial time point for constructing past vector):

$$Y_p = [\boldsymbol{y}_p(p), \boldsymbol{y}_p(p+1), \cdots, \boldsymbol{y}_p(p+m-1)]$$
$$Y_f = [\boldsymbol{y}_f(p), \boldsymbol{y}_f(p+1), \cdots, \boldsymbol{y}_f(p+m-1)]. \qquad (3)$$

The time lags, p and f, can be estimated by calculating the auto-correlation of time-lagged variable vectors. Analogically to PCA, the quasi-covariance matrix H is defined by covariance and cross-covariance matrices of Y_p and Y_f:

$$\Sigma_{pp} = Y_p^{\mathrm{T}} Y_p, \quad \Sigma_{ff} = Y_f^{\mathrm{T}} Y_f, \quad \Sigma_{fp} = Y_f^{\mathrm{T}} Y_p \qquad (4)$$

$$H = \Sigma_{ff}^{-\frac{1}{2}} \Sigma_{fp} \Sigma_{pp}^{-\frac{1}{2}}. \qquad (5)$$

Consequentially, the projection matrices J and L are the normalized results of singular value decomposition result of H:

$$H = U \Lambda V^{\mathrm{T}} \qquad (6)$$

$$J = V_d \Sigma_{pp}^{-\frac{1}{2}}, \quad L = (I_r - V_d V_d^{\mathrm{T}}) \Sigma_{pp}^{-\frac{1}{2}}. \qquad (7)$$

The canonical variate vector $\boldsymbol{z}$ and residual vector $\boldsymbol{e}$ are both linear projections of past vector $\boldsymbol{y}_p$ at time t:

$$\boldsymbol{z}(t) = J\boldsymbol{y}_p(t), \quad \boldsymbol{e}(t) = L\boldsymbol{y}_p(t). \qquad (8)$$

2.2 State-space ICA

Another example of multi-layer linear feature extraction algorithms is the state-space ICA[13]. Following the algorithm structure in Fig. 1, CVA comes as the first layer for handling process dynamics and ICA is conducted in succession to both canonical variate space and residual space to account for non-Gaussianity. These independent components are used for monitoring statistics calculation.

In previous section, (8) have been derived for calculating canonical variates and residuals in CVA approach. Further linear projection of canonical variate space and residual space in order to maximize the independence

between features are defined in (9) in state-space ICA:

$$\boldsymbol{s}(t) = W_z \boldsymbol{z}^{\mathrm{T}}(t), \quad \boldsymbol{q}(t) = W_e \boldsymbol{e}^{\mathrm{T}}(t). \qquad (9)$$

The explicit solution to the projection matrices W_z and W_e, such that the independence of individual components is maximized, is not available; instead, iterative optimization approaches like FastICA[25] may be used.

2.3 PCA-enhanced state-space ICA

The development of PCA-enhanced CVA and state-space ICA indicates that applying multiple feature extraction layers successively may improve the performance of monitoring algorithms. Therefore, PCA-enhanced state-space ICA with 3 layers of feature extraction is proposed as a further extension:

1) Apply PCA to the original data set for dimension-reduced principal components.

2) Apply CVA to the Henkel matrix constructed by principal components in order to handle process dynamics.

3) Apply ICA to the canonical variates and residuals for extraction of representative features.

The structure of PCA-enhanced state-space ICA algorithm can also be summarized using Fig. 1.

2.4 Fault detection stage

After features have been extracted from original data, monitoring statistics are to be calculated using these features and compared with their control limits for fault detection. Qin[26] has studied a variety of monitoring metrics in data-driven process monitoring. The most widely used ones among all are the T^2 statistics for detecting systematic variation using representative features (e.g., canonical variates in CVA and independent components in ICA) and Q statistics for random error using residuals, shown in (10).

$$T^2(t) = \boldsymbol{z}^{\mathrm{T}}(t)\boldsymbol{z}(t), \quad Q(t) = \boldsymbol{e}^{\mathrm{T}}(t)\boldsymbol{e}(t). \qquad (10)$$

Based on normal data, upper control limits with confidence level α, i.e., $T^2_{UCL}(\alpha)$ and $Q_{UCL}(\alpha)$, of statistics in (10) are defined as

$$P(T^2 > T^2_{UCL}(\alpha)) = \alpha, \quad P(Q > Q_{UCL}(\alpha)) = \alpha. \qquad (11)$$

Due to the potential non-Gaussianity of the process variables, the distribution functions in (11) and corresponding control limits are estimated via kernel density estimation[27]. In online fault detection, monitoring statistics calculated with the real-time measurements are compared with these control limits to determine the fault occurrence based on the condition shown in (12).

$$\left(T^2(t) > T^2_{UCL}\right) || \left(Q(t) > Q_{UCL}\right). \qquad (12)$$

3 Contribution propagation analysis

This section discusses the general formulation of contribution plots under multi-layer feature extraction framework and derives the solution to PCA-enhanced, state-space ICA, and PCA-enhanced state-space ICA methods.

3.1 Contribution propagation

Equation (13) formulates the structure of a general 2-layer linear feature extraction based fault detection method.

$$\begin{aligned} \boldsymbol{y} &= P\boldsymbol{x} \\ \boldsymbol{z} &= J\boldsymbol{y}, \quad \boldsymbol{e} = L\boldsymbol{y} \\ T^2 &= \boldsymbol{z}^{\mathrm{T}}\boldsymbol{z}, \quad Q = \boldsymbol{e}^{\mathrm{T}}\boldsymbol{e} \end{aligned} \tag{13}$$

where

$\boldsymbol{x}$: original measured variable vector;

P: first layer projection matrix;

$\boldsymbol{y}$: intermediate features from first layer;

J: second layer projection matrix for feature variables representing systematic error;

L: second layer projection matrix for residuals representing random error;

$\boldsymbol{z}, \boldsymbol{e}$: features from second layer representing systematic and random error;

T^2, Q: monitoring statistics for systematic and random error.

In this formulation, the original data set $X = [\boldsymbol{x}(1), \boldsymbol{x}(2), \cdots, \boldsymbol{x}(n)]^{\mathrm{T}}$ is initially processed by the first feature extraction method to obtain the intermediate feature data $Y = [\boldsymbol{y}(1), \boldsymbol{y}(2), \cdots, \boldsymbol{y}(n)]^{\mathrm{T}}$. Y is further processed by the second layer of feature extraction method to attain the feature variables $Z = [\boldsymbol{z}(1), \boldsymbol{z}(2), \cdots, \boldsymbol{z}(n)]^{\mathrm{T}}$ and residual variables $E = [\boldsymbol{e}(1), \boldsymbol{e}(2), \cdots, \boldsymbol{e}(n)]^{\mathrm{T}}$. The final monitoring statistics are based on $\boldsymbol{z}$ and $\boldsymbol{e}$. It is obvious that aforementioned PCA-enhanced CVA and State-space ICA fall into this category.

The objective of contribution plots-based fault diagnosis is to gain the contribution of original process variables to the final monitoring statistics such as T^2 and Q. In order to do so, Fig. 2 illustrates the propagation of variable contributions under this multi-layer feature extraction framework.

When the number of intermediate layers increases, the propagation procedure of contributions will remain the same. Hence, this framework also applies to PCA-enhance State-space ICA algorithm which has three layers of feature extraction.

The following equations hold for individual variables $x_i \in \boldsymbol{x}$, $y_j \in \boldsymbol{y}$ and $z_k \in \boldsymbol{z}$: $\sum_{i=1}^{v} cont_{x_i, y_j} = y_j$; $\sum_{j=1}^{r} cont_{y_j, z_k} = z_k$; $cont_{z_k, T^2} = z_k^{\mathrm{T}} z_k$. Analogy can be made for the contribution plots to Q statistics. The general philosophy behind is to calculate the weighted com-
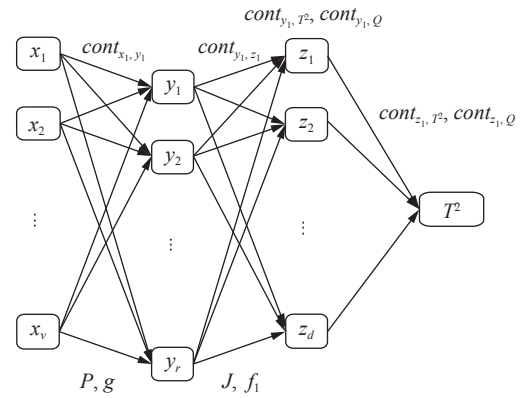

Fig. 2 Illustration of 2-layer contribution propagation in T^2

bination of the contributions of intermediate features to the final statistics ($cont_{y_j, T^2}$ and $cont_{y_j, Q}$), in which the weighting coefficients are the contribution of original variable to the intermediate features ($cont_{x_i, y_j}$).

In general, the individual contributions of single variables are displayed in (14), where $P_{i,j}$ and $J_{j,k}$ are intermediate coefficients with respect to y_j and z_k, respectively.

$$\begin{aligned} cont_{x_i, y_j} &= x_i P_{i,j} \\ cont_{y_j, z_k} &= y_j J_{j,k} \\ cont_{x_i, z_k} &= \sum_{j=1}^{r} x_i J_{j,k} P_{i,j}. \end{aligned} \tag{14}$$

PCA, CVA and ICA are all linear layers; therefore, a close form solution to the complete contribution plots of original variables to monitoring statistics can be derived explicitly.

3.2 Contribution calculation for PCA-enhanced CVA-based monitoring

Following the general formulation proposed previously, the contribution plots of this PCA-enhanced CVA algorithm can be derived in 4 steps.

1) Contribution of the original process variable vector $\boldsymbol{x}$ to the principal components $\boldsymbol{y}$.

The general process model of PCA is

$$\boldsymbol{y}^{\mathrm{T}} = \boldsymbol{x}^{\mathrm{T}} P = \sum_{i=1}^{v} x_i \boldsymbol{p}_i \tag{15}$$

where $\boldsymbol{x} = [x_1, x_2, \cdots, x_v]^{\mathrm{T}}$ and x_i is the i-th process variable, whilst $\boldsymbol{p}_i \in \mathbf{R}^r$ is the i-th row vector of P. Therefore, for each variable x_i, its contribution to the entire principal components $\boldsymbol{y}$ is calculated by (16):

$$cont_{x_i, \boldsymbol{y}}(t) = x_i(t) \boldsymbol{p}_i. \tag{16}$$

2) Contribution of principal components $\boldsymbol{y}$ to canonic-

al variates $\boldsymbol{z}$ and residuals $\boldsymbol{e}$: Based on the original definition of contribution, the contributions of the j-th principal component y_j to the k-th canonical variate z_k and residuals e_h at time t are calculated by decomposing z_k and e_h into summation of individual contributions.

$$cont_{y_j,z_k}(t) = \sum_{l=1}^{p} |J_{j_l,k} y_j(t-l)|$$
$$cont_{y_j,e_h}(t) = \sum_{l=1}^{p} |L_{j_l,h} y_j(t-l)| \qquad (17)$$

where $J_{j_l,k}$ is the coefficient of the j-th principal component y_j with time delay l with respect to the k-th canonical variate z_k. J and L are the projection matrices obtained by CVA.

3) Contribution of canonical variates $\boldsymbol{z}$ and residuals $\boldsymbol{e}$ to monitoring statistics: According to (10), the monitoring statistics can be decomposed into summations of squared terms of individual features, as shown in (18).

$$cont_{z_k,T^2}(t) = z_k^2(t)$$
$$cont_{e_h,Q}(t) = e_h^2(t). \qquad (18)$$

4) Contribution of the i-th process variables x_i to monitoring statistics: Noticing that based on (16), the contribution of x_i to y_j at time $t-l$ is $cont_{x_i,y_j}(t-l) = x_i(t-l)p_{i,j}$, the contribution of variable x_i to T^2 and Q at time t can be obtained by (19) and (20).

$$cont_{x_i,\mathrm{T^2}}(t) = \sum_{j=1}^{d} \sum_{l=1}^{p} \left| \boldsymbol{z}^{\mathrm{T}}(t) J_{j_l} cont_{x_i,y_j}(t-l) \right| = \left| \boldsymbol{z}^{\mathrm{T}}(t) \right| |J| \left| \boldsymbol{c}_{i,p}(t) \right| \qquad (19)$$

$$cont_{x_i,T^2}(t) = \sum_{j=1}^{d} \sum_{l=1}^{p} \left| \boldsymbol{z}^{\mathrm{T}}(t) J_{j_l} cont_{x_i,y_j}(t-l) \right| = \left| \boldsymbol{z}^{\mathrm{T}}(t) \right| |J| \left| \boldsymbol{c}_{i,p}(t) \right| \qquad (20)$$

where the $\boldsymbol{c}_{i,p}(t) \in \mathbf{R}^{pf}$ is the t-th column vector of past Hankel matrix $C_{i,p}$ of $x_i\boldsymbol{p}_i$ which is constructed the same way as Y_p.

3.3 Contribution calculation for state-space ICA based monitoring

Similarly, the contribution propagation for state-space ICA can be obtained:

1) Contribution of the original process variable vector $\boldsymbol{x}$ to the canonical variates $\boldsymbol{z}$ and residuals $\boldsymbol{e}$: Similarly to the second step in PCA-enhanced CVA, (21) defines the contribution of the i-th process variable to the k-th canonical variate z_k and the h-th residual e_h:

$$cont_{x_i,z_k}(t) = \sum_{l=1}^{p} |J_{i_l,k} x_i(t-l)|$$
$$cont_{x_i,e_h}(t) = \sum_{l=1}^{p} |L_{i_l,h} x_i(t-l)|. \qquad (21)$$

2) Contribution of $\boldsymbol{z}$ and $\boldsymbol{e}$ to independent components $\boldsymbol{s}$ and $\boldsymbol{r}$: Since both CVA and ICA are linear projections of original features, the contribution calculation resembles the procedure in (16).

$$cont_{z_k,\boldsymbol{s}}(t) = z_k(t)\boldsymbol{w}_{z,k}$$
$$cont_{e_h,\boldsymbol{r}}(t) = e_h(t)\boldsymbol{w}_{r,h} \qquad (22)$$

where $\boldsymbol{w}_{z,k}$ and $\boldsymbol{w}_{r,h}$ are column vectors of W_z and W_e.

3) Contribution of $\boldsymbol{s}$ and $\boldsymbol{r}$ to monitoring statistics: The contribution of s_α, α-th independent component in canonical space, to T^2 and the contribution of r_β to Q are

$$cont_{s_\alpha,T^2}(t) = s_\alpha^2(t)$$
$$cont_{r_\beta,Q}(t) = r_\beta^2(t). \qquad (23)$$

4) Contribution of original variables to monitoring statistics: The contribution of variable x_i to T^2 and Q at time t can be obtained by (24) and (25).

$$cont_{x_i,T^2}(t) = \boldsymbol{s}^{\mathrm{T}}(t) \sum_{\alpha} \sum_{l=1}^{p} |\boldsymbol{w}_{z;\alpha} J_{i_l} x_i(t-l)| \qquad (24)$$

$$cont_{x_i,Q}(t) = \boldsymbol{r}^{\mathrm{T}}(t) \sum_{\beta} \sum_{l=1}^{p} |\boldsymbol{w}_{e;\beta} L_{i_l} x_i(t-l)| \qquad (25)$$

where $\boldsymbol{w}_{z;\alpha}$ and $\boldsymbol{w}_{e;\beta}$ are row vectors of W_z and W_e.

Note that due to the mathematical formulation of CVA and ICA, the covariance matrix of extracted features is diagonal. Therefore, the cross terms of canonical variates/independent components do not exist in monitoring statistics and these statistics can be easily decomposed into summation of contributions of individual features. In other linear algorithms of which the covariance matrix is not diagonal, the contribution propagation framework will result in an approximated solution.

3.4 Contribution calculation for PCA-enhanced state-space ICA based monitoring

Similarly to PCA-enhanced CVA and state-space ICA, contribution propagation analysis calculates the contributions propagated in each layer and estimates the contribution of original variables to final monitoring statistics by synthesizing the propagated contributions when PCA-enhanced state-space ICA is adopted. To be brief, the procedure is summarized as follows:

1) Contribution of original process variables to prin-

cipal components;

2) Contribution of principal components to canonical variates and residuals;

3) Contribution of canonical variates and residuals to corresponding independent components;

4) Contribution of independent components to monitoring statistics;

5) Contribution of original process variables to monitoring statistics.

3.5 Algorithm summary

To summarize, the online fault detection and diagnosis procedure based on multi-layer linear feature extraction techniques is follows:

1) Retrieve the new sample vector $\boldsymbol{x}^*$ and calculate the intermediate features $\boldsymbol{y}^*$ with the first layer of feature extraction algorithm;

2) Continue the calculation until the monitoring features $\boldsymbol{z}^*$ and residuals $\boldsymbol{e}^*$ are obtained;

3) Calculate the monitoring statistics using $\boldsymbol{z}^*$ and $\boldsymbol{e}^*$;

4) If the monitoring statistics exceed their control limits, calculate the contribution of all x_i^* to the monitoring statistics propagated through multiple layers so as to identify the influential variables.

4 Case studies

In this section, the proposed contribution propagation framework will be validated using two case studies: a simulated benchmark case study of the Tennessee Eastman process and a benchmark data set collected by experiments on a multiphase flow facility.

4.1 TE plant

4.1.1 Process description

As a simulated case study, the Tennessee Eastman process (TEP) has been widely acknowledged as a bench-

mark for development and validation of process monitoring and control algorithms. The process configuration of TEP used in this work is inherited from [28]. Simulated data set consisting of 22 measured variables is used for training and validation; the tag names and descriptions of these variables are presented in Table 1.

To validate, a training set of 1 000 samples is collected in normal operating condition. For validation, two extra data sets of 1 000 samples each are generated under in presence of two fault scenarios (IDV(11) and IDV(13) in [28]), respectively. The fault descriptions are provided in Table 2. The faulty period starts at Sample 200 and ends at 1 000.

4.1.2 Fault diagnosis results

It has been demonstrated in previous works that PCA-enhanced CVA and state-space ICA can improve fault detection performance on the TEP data set by comparing with CVA and dynamic extensions of PCA and ICA[12, 13]. Therefore, this section focuses on the fault diagnosis performance obtained by contribution propagation framework used in multi-layer linear feature extractions and the ordinary contribution plots applied to CVA-based monitoring.

The contribution plots of all 4 monitoring algorithms for both faults are presented in Figs. 3 and 4 in the order of the number of layers increases.

Table 3 further summarizes the influential variables identified by their contributions in CVA, PCA-enhanced CVA, state-space ICA, and PCA-enhanced state-space ICA. The union set of variables that rank the top 2 in their contributions to T^2 and Q statistics is identified as influential variables. A more systematic way of identifying influential variables using the control limits of variable contributions[21] can be considered as a potential extension.

For IDV(11), it has been agreed by the contribution plots of all 4 algorithms that XMEAS(9), the reactor temperature, is the most directly influenced variable when the reactor cooling water inlet has large random

Table 1   Process variables

| Number	Variable name	Number	Variable name
XMEAS (1)	A feed (stream 1)	XMEAS (12)	Separator level
XMEAS (2)	D feed (stream 2)	XMEAS (13)	Separator pressure
XMEAS (3)	E feed (stream 3)	XMEAS (14)	Separator underflow
XMEAS (4)	A and C feed	XMEAS (15)	Stripper level
XMEAS (5)	Recycle flow	XMEAS (16)	Stripper pressure
XMEAS (6)	Reactor feed rate	XMEAS (17)	Stripper underflow
XMEAS (7)	Reactor pressure	XMEAS (18)	Stripper temperature
XMEAS (8)	Reactor level	XMEAS (19)	Stripper steam flow
XMEAS (9)	Reactor temperature	XMEAS (20)	Compressor work
XMEAS (10)	Purge rate	XMEAS (21)	Reactor water temperature
XMEAS (11)	Separator temperature	XMEAS (22)	Separator water temperature
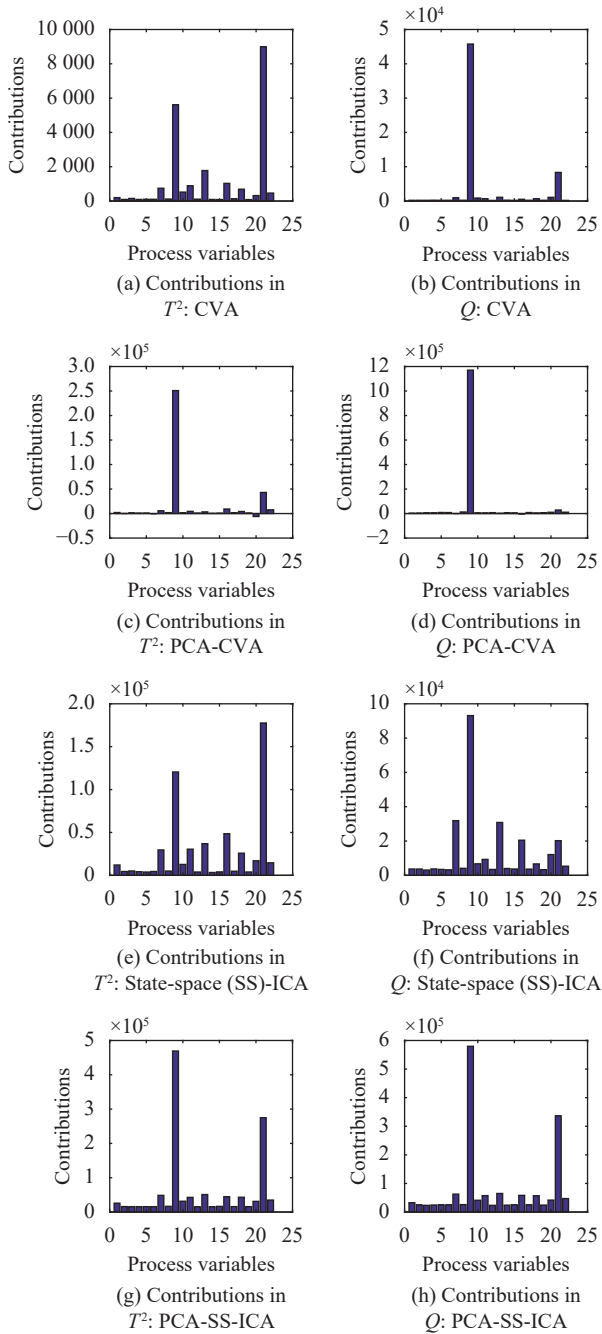
Fig. 3   Contribution performance: IDV(11)

Table 2   Fault specification in TEP case study

	Process variable	Type
IDV(11)	Reactor cooling water inlet	Random variation
IDV(13)	Reaction kinetics	Slow drift

variation; moreover, the reactor water temperature (XMEAS(21)) will also be influenced. This conclusion can be justified by reasoning with respect to the process configuration: Since the reactions in this plant are all exothermic[29], the large variations of coolant will cause vari-
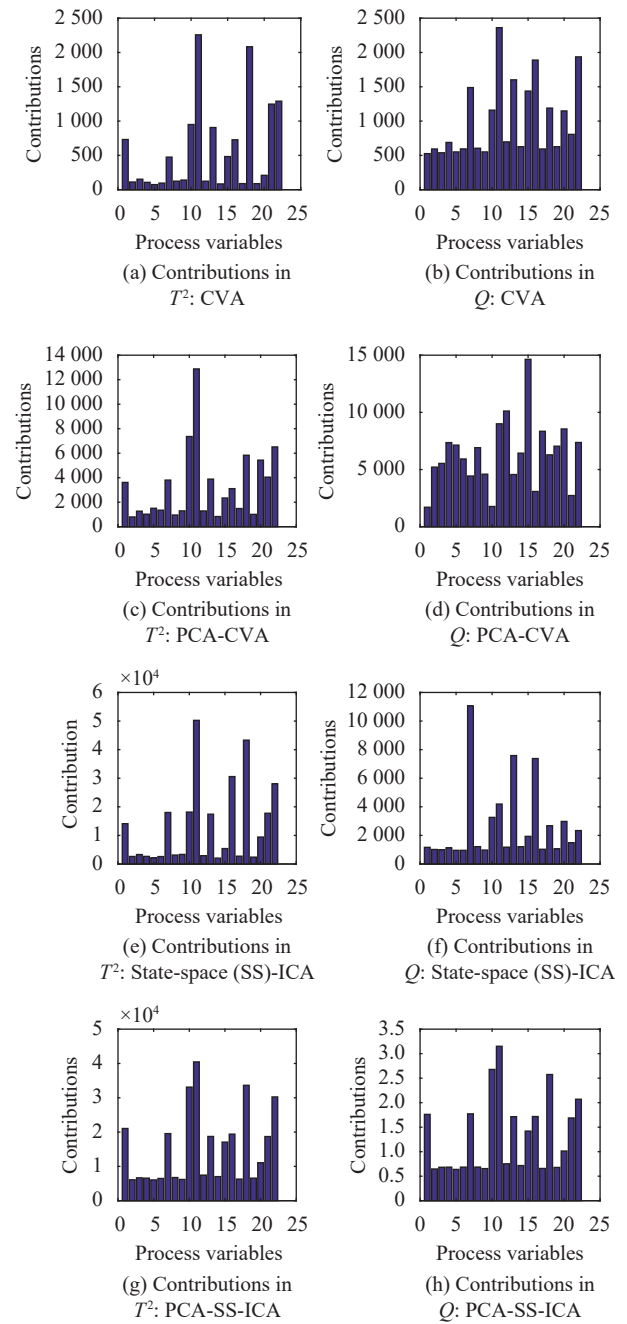
Fig. 4   Contribution performance: IDV(13)

ations in the temperature measurements associated to the reactor; large variations in XMEAS(9) and XMEAS(21), therefore, are the cause of monitoring statistics exceeding their thresholds. This observation can further assist locating the fault.

IDV(13) is more difficult to diagnose since it is hard to quantify the reaction kinetics directly. Since the reactants are all gas and resultants are all liquids in reactions of this plant[29], the pressure drop in the system will be a good indicator of reaction kinetics, which is reflected in reactor pressure (XMEAS(7)). Since after the condenser, gas resultants will liquefy and the pressure is no longer

Table 3 Fault isolation

	IDV(11)
CVA	XMEAS (9), XMEAS (21)
PCA-CVA	XMEAS (9)
SS-ICA	XMEAS (9), XMEAS (21)
PCA-SS-ICA	XMEAS (9), XMEAS (21)
	IDV(13)
CVA	XMEAS (11), XMEAS (18), XMEAS (22)
PCA-CVA	XMEAS (11), XMEAS (15)
SS-ICA	XMEAS (7), XMEAS (11), XMEAS (18)
PCA-SS-ICA	XMEAS (10), XMEAS (11), XMEAS (18)

dominant in the separator or stripper, the pressure variation in the stripper will be less significant. Instead, the stripper level (XMEAS(15)) will be influenced owing to the insufficient reaction. Variations in temperature also propagates from reactor (XMEAS(11)) to the stripper (XMEAS(18)) due to the exothermic nature of this reaction. These variables indicate that this fault impacts the condition in reactor and other components downstream and therefore restrain the area of fault existence. Fig. 5 visualizes the trend plot influential variables identified by state-space ICA (XMEAS(7) and XMEAS(11)) and a clear trend of variation can be observed after the fault was seeded.
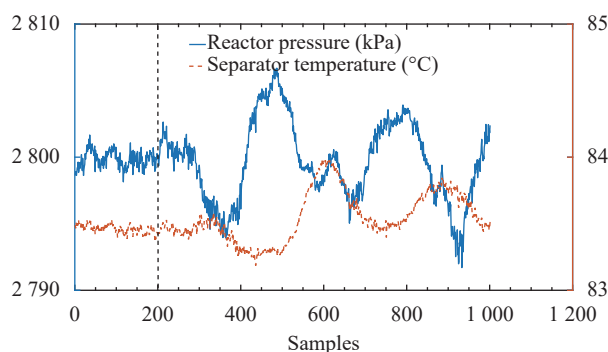

Fig. 5 Trend plot of influential variables

It can be concluded that, in this simulated case study, the proposed contribution propagation framework provides a fault isolation approach for PCA-enhanced CVA, state-space ICA, and PCA-enhanced state-space ICA and identifies influential variables. While the fault detection performance is improved by multiple layers of feature extraction, the contribution of original measured variables can still be estimated via contribution propagation and hence facilitates fault diagnosis afterwards.

4.2 Multiphase flow benchmark

4.2.1 Process description

The multiphase flow facility in the Process System Engineering Lab of Cranfield University is a unique industrial-scale rig for researches and experiments on measuring, monitoring and control of multiphase flows. Water, oil and air are supplied from individual pipelines; by converging and intersection of pipelines, 3-phase flows are mixed, making a multiphase flow with liquid and gas. The multiphase flow is transported, measured, separated and recycled successively afterwards. Being fully automated, this facility can operate in multiple normal operating conditions as well as simulate various faulty scenarios with manually seeded faults. It is also well equipped with measurement instrumentations which contain both regular process variables such as pressure and temperature, and mechanical condition variable such as pump current. All measurement data are collected in real-time and recorded by DeltaV system for further analysis. A more detailed description of this benchmark case study and previous work on statistical monitoring of it can be found in [4] and [30].

The schematic with the layout of measurement instrumentations of this facility is shown in Fig. 6. A total of 23 process variables are measured and recorded in the benchmark data set and variable descriptions are provided by Table 4.

4.2.2 Fault diagnosis results

In the preliminary fault detection step, CVA, PCA-enhanced CVA, state-space ICA, and PCA-enhanced state-space ICA algorithms are used for calculating the monitoring statistics separately; the propagated contributions of original variables to final monitoring statistics in multilayer algorithms are calculated with respect to Section 3. The results obtained from proposed framework are compared with those obtained directly by CVA-based monitoring using original contribution plots[4]. Fig. 7 compares the contribution plots of all variables obtained by CVA, PCA-enhanced CVA, state-space ICA, and PCA-enhanced state-space ICA in Fault 1. Their results in Fault 2 are compared in Fig. 8. To summarize the fault diagnosis results suggested by contributions, Table 5 shows the influential variables identified by contributions of variables using different monitoring algorithms.

For Fault 1, pipeline blockage is mimicked by turning the control valve on the input pipeline to top riser (VC404) gradually. According to the contribution plots, PCA-enhanced CVA and PCA-enhanced state-space ICA suggest the differential pressure over VC404 (V7) is the most influential variable while sate-space ICA and CVA both identify the riser top pressure (V3) as an influential variable too. Noticing that the valve opening of VC404 is not involved as process variable, the pressure drop over this valve and pressure measurements in adjacent to it will be the proper indicator of pipeline blockage. Therefore, the influential variables, which the deviation and original contribution plots both agree upon, are the most relevant ones among all with respect to Fault 1 and the results of contribution-based fault diagnosis are convincing.
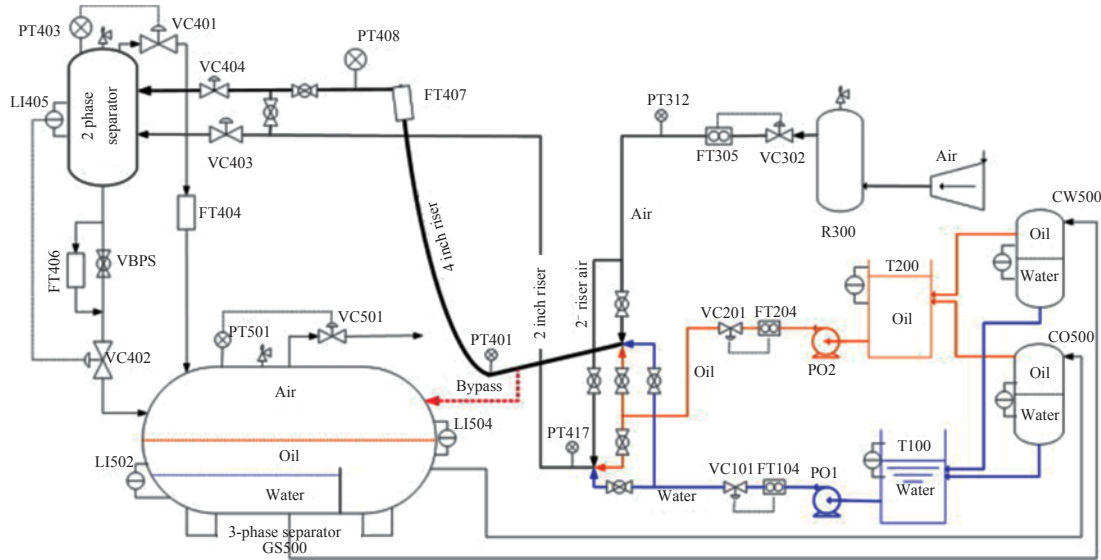
Fig. 6   Schematic of the multiphase flow facility

Table 4   Measured variables in multiphase flow facility

Number	Description	Location	Number	Description	Location
1	Air delivery pressure	PT312	13	Top riser density	FT407
2	Riser bottom pressure	PT401	14	Top separator output density	FT406
3	Riser top pressure	PT408	15	Input water density	FT104
4	Top separator pressure	PT403	16	Top riser temperature	FT407
5	3 phase separator pressure	PT501	17	Top separator output temperature	FT406
6	Differential pressure (PT401-PT408)	PT408	18	Input water temperature	FT104
7	Differential pressure over VC404	PT403	19	3 phase separator gas-liquid level	LI504
8	Input air flow rate	FT305	20	Valve position of VC501	VC501
9	Input water flow rate	FT104	21	Valve position of VC302	VC302
10	Top riser flow rate	PT403	22	Valve position of VC101	VC101
11	Top separator level	LI405	23	Water pump current	PO1
12	Top separator output flow rate	FT406			

For Fault 2, due to the insufficient air/water supply, liquid accumulates at the bottom of the riser and blocks the gas transportation, causing the bottom pressure to increase. The liquid level will continue rising until blocked gas penetrates and results in liquid blow out at the riser top. Remaining liquid will fall back to the riser bottom and accumulate again. From process data perspective, the slugging fault will cause large fluctuations in process variables, such as pressure, flow rate and density, at both riser top and riser bottom. However, unlike single variable faulty scenario, the identified influential variables from different algorithms in slugging condition are no longer unified. For instance, the differential pressure between riser bottom and top (V6), riser bottom pressure (V2), and riser top density (V13) are identified commonly while variables in vicinity, such as riser top flow rate (V10) and riser top pressure (V3), are also suggested by different algorithms. Roughly speaking, all vari-

ables with large contributions are in line with the fact that variables at riser top and bottom will be mostly influenced when slugging occurs; while the contribution of specific measurements will be determined by the corresponding feature extraction algorithm applied.

It can be concluded that the proposed contribution propagation framework is capable of estimating the contribution of original variables in monitoring statistics when multi-layer linear feature extraction method is employed in monitoring statistics calculation; therefore it is suitable for fault diagnosis and isolation of relevant variables in these situations. The fault diagnosis performance can be maintained as the fault detection performance is improved by substituting simple feature extraction methods with multi-layer feature ones. In the meantime, variations may exist in diagnosis results due to different feature extraction methods applied, especially in presence of complicated fault that influences multiple process vari-
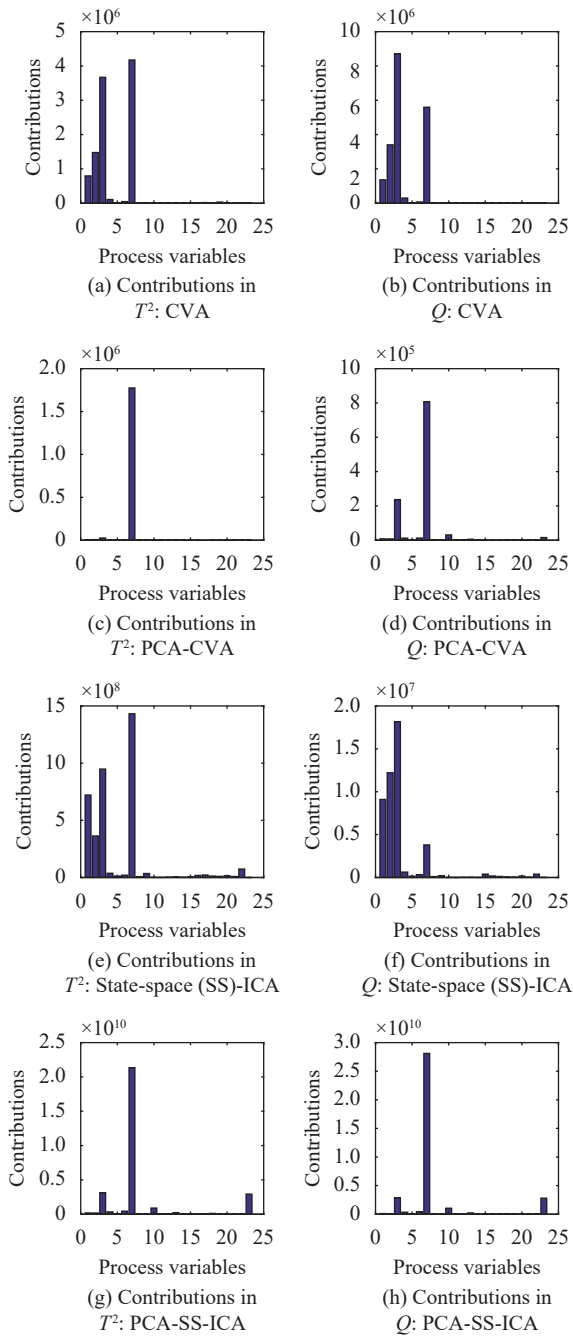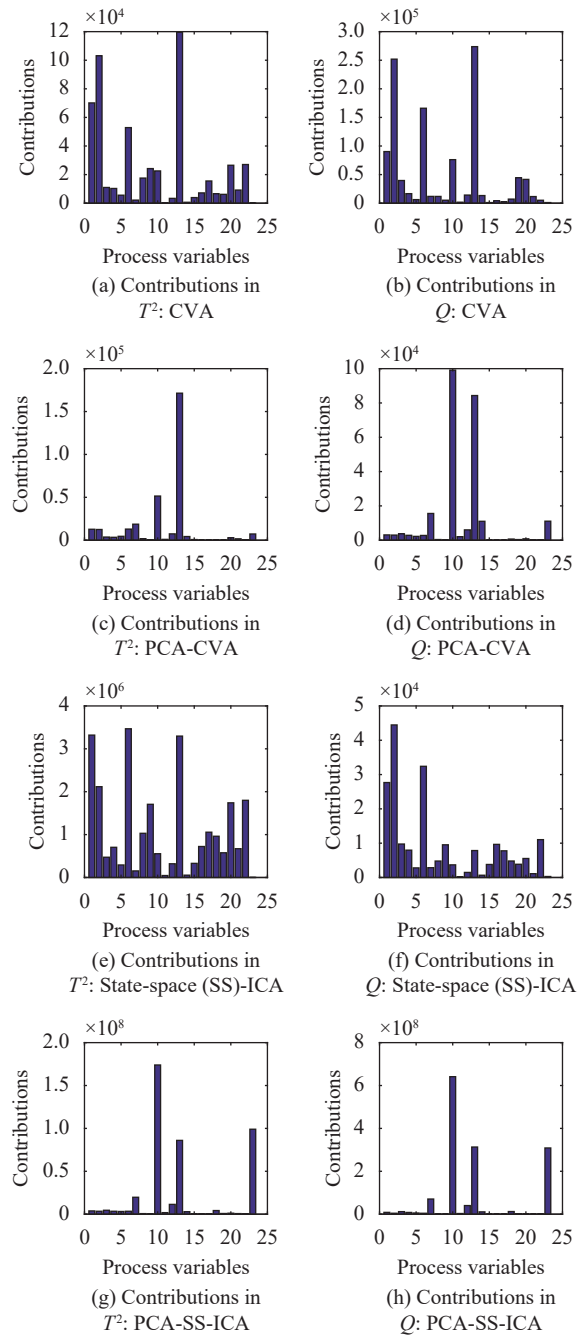
Fig. 7   Contribution performances: Fault 1

Table 5   Influential variables identified

	Fault 1	Fault 2
CVA	V3, V7	V3, V2, V6
PCA-CVA	V7	V10, V13
SS-ICA	V3, V7	V6, V2, V13
PCA-SS-ICA	V7	V10, V13

ables, as observed from this case study.

4.2.3 Discussions

The case study further demonstrates the fault isola-

Fig. 8   Contribution performance: Fault 2

tion ability of proposed contribution propagation approach for PCA-enhanced CVA and state-space ICA in data set collected from real process. Moreover, the solution can be visualized and interpreted the same way as the ordinary contribution plots in basic feature extraction approaches. On the other hand, there still exists potential of improvement for detection and diagnosis of the slugging fault.

5 Conclusions and future work

In this work, we proposed contribution propagation

analysis, which is a novel fault diagnosis approach based on the original contribution concept, for fault diagnosis and influential process variable identification that applies when multi-layer linear feature extraction algorithms are used for fault detection. As a more advanced example of multi-layer linear algorithms, PCA-enhanced state-space ICA has also been proposed in addition to existing PCA-enhanced CVA and state-space ICA. By the validation of the TE data and the multiphase flow facility data, it has been demonstrated that this contribution propagation analysis is compatible with various types of multi-layer feature extraction algorithms and the interpretability of the monitoring results obtained by these existing algorithms has been improved significantly.

The gap of contribution plots-based fault diagnosis still exists in this multi-layer contribution propagation framework for monitoring of processes and faults with extra complexity due to the mathematical complexity of nonlinear approaches adopted, such as kernel transformation. Therefore, it is worthwhile considering the propagation of contribution plots in context of "kernelized" and other advanced feature extraction methods so as to provide a general solution to the contribution plots of variables for different monitoring techniques in the future.

Acknowledgements

Open Access

References

[1] S. J. Ren, F. Q. Si, J. X. Zhou, Z. L. Qiao, Y. L. Cheng. A new reconstruction-based auto-associative neural network for fault diagnosis in nonlinear systems. *Chemometrics and Intelligent Laboratory Systems*, vol. 172, pp. 118–128, 2018. DOI: 10.1016/j.chemolab.2017.12.005.

[2] H. Ren, Y. Chai, J. F. Qu, X. Ye, Q. Tang. A novel adaptive fault detection methodology for complex system using deep belief networks and multiple models: A case study on cryogenic propellant loading system. *Neurocomputing*, vol. 275, pp. 2111–2125, 2018. DOI: 10.1016/j.neucom.2017.10.063.

[3] J. E. Jackson, G. S. Mudholkar. Control procedures for residuals associated with principal component analysis. *Technometrics*, vol. 21, no. 3, pp. 341–349, 1979. DOI: 10.1080/00401706.1979.10489779.

[4] C. Ruiz-Cárcel, Y. Cao, D. Mba, L. Lao, R. T. Samuel. Statistical process monitoring of a multiphase flow facility. *Control Engineering Practice*, vol. 42, pp. 74–88, 2015. DOI: 10.1016/j.conengprac.2015.04.012.

[5] L. H. Chiang, E. L. Russell, R. D. Braatz. Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, vol. 50, no. 2, pp. 243–252, 2000. DOI: 10.1016/S0169-7439(99)00061-1.

[6] C. H. Zhao, F. L. Wang, Z. H. Mao, N. Y. Lu, M. X. Jia. Adaptive monitoring based on independent component analysis for multiphase batch processes with limited modeling data. *Industrial & Engineering Chemistry Research*, vol. 47, no. 9, pp. 3104–3113, 2008. DOI: 10.1021/ie701680y.

[7] M. Misra, H. H. Yue, S. J. Qin, C. Ling. Multivariate process monitoring and fault diagnosis by multi-scale PCA. *Computers & Chemical Engineering*, vol. 26, no. 9, pp. 1281–1293, 2002. DOI: 10.1016/S0098-1354(02)00093-5.

[8] M. F. Harkat, S. Djelel, N. Doghmane, M. Benouaret. Sensor fault detection, isolation and reconstruction using nonlinear principal component analysis. *International Journal of Automation and Computing*, vol. 4, no. 2, pp. 149–155, 2007. DOI: 10.1007/s11633-007-0149-6.

[9] B. B. Jiang, X. X. Zhu, D. X. Huang, J. A. Paulson, R. D. Braatz. A combined canonical variate analysis and Fisher discriminant analysis (CVA-FDA) approach for fault diagnosis. *Computers & Chemical Engineering*, vol. 77, pp. 1–9, 2015. DOI: 10.1016/j.compchemeng.2015.03.001.

[10] R. T. Samuel, Y. Cao. Kernel canonical variate analysis for nonlinear dynamic process monitoring. *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 605–610, 2015. DOI: 10.1016/j.ifacol.2015.09.034.

[11] Z. Q. Ge, S. Y. Zhong, Y. W. Zhang. Semisupervised kernel learning for FDA model and its application for fault classification in industrial processes. *IEEE Transactions on Industrial Informatics*, vol. 12, no. 4, pp. 1403–1411, 2016. DOI: 10.1109/TII.2016.2571680.

[12] Y. Cao, R. T. Samuel. Dynamic latent variable modelling and fault detection of Tennessee Eastman challenge process. In *Proceedings of IEEE International Conference on Industrial Technology*, IEEE, Taipei, China pp. 842–847, 2016. DOI: 10.1109/ICIT.2016.7474861.

[13] P. P. Odiowei, Y. Cao. State-space independent component analysis for nonlinear dynamic process monitoring. *Chemometrics and Intelligent Laboratory Systems*, vol. 103, no. 1, pp. 59–65, 2010. DOI: 10.1016/j.chemolab.2010.05.014.

[14] Z. Q. Ge, Z. H. Song. Process monitoring based on independent component analysis-principal component analysis (ICA-PCA) and similarity factors. *Industrial & Engineering Chemistry Research*, vol. 46, no. 7, pp. 2054–2063, 2007. DOI: 10.1021/ie061083g.

[15] C. H. Zhao, F. R. Gao, F. L. Wang. Nonlinear batch process monitoring using phase-based kernel-independent component analysis-principal component analysis (KICA-PCA). *Industrial & Engineering Chemistry Research*, vol. 48, no. 20, pp. 9163–9174, 2009. DOI: 10.1021/ie8012874.

[16] H. K. Chen, X. G. Zhao, S. Y. Sun, M. Tan. PLS-CCA heterogeneous features fusion-based low-resolution human detection method for outdoor video surveillance. *International Journal of Automation and Computing*, vol. 14, no. 2, pp. 136–146, 2017. DOI: 10.1007/s11633-016-1029-8.

[17] P. Miller, R. E. Swanson, C. E. Heckler. Contribution plots: A missing link in multivariate quality control. *Applied Mathematics and Computer Science*, vol. 8, no. 4, pp. 775–792, 1998.

[18] B. B. Jiang, D. X. Huang, X. X. Zhu, F. Yang, R. D. Braatz. Canonical variate analysis-based contributions for fault identification. *Journal of Process Control*, vol. 26, pp. 17–25, 2015. DOI: 10.1016/j.jprocont.2014.12.001.

[19] H. Babamoradi, F. van den Berg, Å. Rinnan. Confidence limits for contribution plots in multivariate statistical process control using bootstrap estimates. *Analytica Chimica Acta*, vol. 908, pp. 75–84, 2016. DOI: 10.1016/j.aca.2016.01.002.

[20] S. A. Lawal, J. Zhang. Actuator fault monitoring and fault tolerant control in distillation columns. *International Journal of Automation and Computing*, vol. 14, no. 1, pp. 80–92, 2017. DOI: 10.1007/s11633-016-1037-8.

[21] J. A. Westerhuis, S. P. Gurden, A. K. Smilde. Generalized contribution plots in multivariate statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems*, vol. 51, no. 1, pp. 95–114, 2000. DOI: 10.1016/S0169-7439(00)00062-9.

[22] C. F. Alcala, S. J. Qin. Reconstruction-based contribution for process monitoring. *Automatica*, vol. 45, no. 7, pp. 1593–1600, 2009. DOI: 10.1016/j.automatica.2009.02.027.

[23] S. W. Choi, J. H. Park, I. B. Lee. Process monitoring using a Gaussian mixture model via principal component analysis and discriminant analysis. *Computers & Chemical Engineering*, vol. 28, no. 8, pp. 1377–1387, 2004. DOI: 10.1016/j.compchemeng.2003.09.031.

[24] R. M. Tan, Y. Cao. Contribution plots based fault diagnosis of a multiphase flow facility with PCA-enhanced canonical variate analysis. In *Proceedings of the 23rd International Conference on Automation and Computing*, IEEE, Huddersfield, UK, 2017. DOI: 10.23919/IConAC.2017.8081992.

[25] A. Hyvärinen, E. Oja. Independent component analysis: Algorithms and applications. *Neural Networks*, vol. 13, no. 4–5, pp. 411–430, 2000. DOI: 10.1016/S0893-6080(00)00026-5.

[26] S. J. Qin. Statistical process monitoring: Basics and beyond. *Journal of Chemometrics*, vol. 17, no. 8–9, pp. 480–502, 2003. DOI: 10.1002/cem.800.

[27] P. P. Odiowei, Y. Cao. Nonlinear dynamic process monitoring using canonical variate analysis and kernel density estimations. *Computer Aided Chemical Engineering*, vol. 27, pp. 1557–1562, 2009. DOI: 10.1016/S1570-7946(09)70650-9.

[28] S. Yin, S. X. Ding, A. Haghani, H. Y. Hao, P. Zhang. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman Process. *Journal of Process Control*, vol. 22, no. 9, pp. 1567–1581, 2012. DOI: 10.1016/j.jprocont.2012.06.009.

[29] J. J. Downs, E. F. Vogel. A plant-wide industrial process control problem. *Computers & Chemical Engineering*, vol. 17, no. 3, pp. 245–255, 1993. DOI: 10.1016/0098-1354(93)80018-I.

[30] Y. Cao. A benchmark case for statistical process monitoring – Cranfield multiphase flow facility, [Online], Available: http://uk.mathworks.com/matlabcentral/fileexchange/50938-a-benchmark-case-for-statistical-process-monitoring-cranfield-multiphase-flow-facility, 2015.
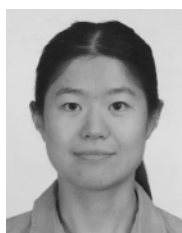
Ruo-Mu Tan received the B. Eng. degree in automation from Zhejiang University, China in 2013, and the M. Sc. degree in process control from the University of Alberta, Canada in 2015. She is currently a Ph. D. degree candidate at chemical engineering at Imperial College London, UK. Meanwhile, she is also a Marie Curie Early Stage researcher of Horizon 2020 Innovative Training Networks – European Industrial Doctorates Project PRONTO (i.e., Process network optimization for efficient and sustainable operation of Europe′s process industries taking machinery condition and process performance into account).

    Her research interests include data-driven nonlinear process monitoring, multivariate statistical analysis and their application to process industries.

    E-mail: r.tan@imperial.ac.uk (Corresponding author)

    ORCID iD: 0000-0002-9151-7245
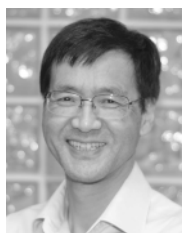
Yi Cao received the M. Sc. degree in control engineering from Zhejiang University, China in 1985 and the Ph. D. degree in engineering from the University of Exeter, UK in 1996. He is a professor in College of Chemical and Biological Engineering, Zhejiang University, China.

    His research interests include plant-wide control, nonlinear system identification, nonlinear model predictive control and process monitoring.

    E-mail: caoyi2018@zju.edu.cn

    ORCID iD: 0000-0003-2360-1485