RESEARCH   **Open Access**

Fault diagnosis of Tennessee Eastman process using signal geometry matching technique

Han Li and De-yun Xiao[*]

Abstract

This article employs adaptive rank-order morphological filter to develop a pattern classification algorithm for fault diagnosis in benchmark chemical process: Tennessee Eastman process. Rank-order filtering possesses desirable properties of dealing with nonlinearities and preserving details in complex processes. Based on these benefits, the proposed algorithm achieves pattern matching through adopting one-dimensional adaptive rank-order morphological filter to process unrecognized signals under supervision of different standard signal patterns. The matching degree is characterized by the evaluation of error between standard signal and filter output signal. Initial parameter settings of the algorithm are subject to random choices and further tuned adaptively to make output approach standard signal as closely as possible. Data fusion technique is also utilized to combine diagnostic results from multiple sources. Different fault types in Tennessee Eastman process are studied to manifest the effectiveness and advantages of the proposed method. The results show that compared with many typical multivariate statistics based methods, the proposed algorithm performs better on the deterministic faults diagnosis.

Keywords: fault diagnosis, pattern matching, adaptive rank-order morphological filtering, Tennessee Eastman process

1. Introduction

The last decades have been witnessing the modern large-scale processes developing toward high complexity and multiplicity in industries such as chemical, metallurgical, mechanical, logistics, and *etc*. These processes are generally characterized by a long-process flow with large operation scales and complicated mechanisms. The typical features are highly nonlinear, long-time delay, and heavily correlated among measurements [1]. Process monitoring, aiming to ensure that the operations satisfy the performance specifications and indicating anomalies, becomes a major challenge in practice. First, the requirements of process expertise for model-based methods often pose difficulties for operators not specializing in this realm; secondly, the system identification theory based methods need to postulate specified mathematical models, which are incapable of capturing varied nonlinearities. In addition, due to the growing number of sensors installed in processes, quantity of data constantly generated under different conditions soars by a few orders of magnitude or more compared to small-

scale processes [2]. The fundamental dilemma for process monitoring is deficient knowledge to establish relative accurate mathematical process description while incomplete methodology to exploit abundant data to reveal process mechanisms and operational statuses. In large-scale processes, standard PI (proportional-integral) or PID (proportional-integral-derivative) closed-loop control schemes are often adopted to compensate for variable disturbances and outliers. However, excessive compensation may easily cause controllers overburden and a trivial glitch could eventually develop to catastrophic fault(s). Based on the considerations of practical limits, demands of safety operation, cost optimization as well as business opportunities in technical development, the problem of how to more effectively utilize mass amount of process data to meet the increasing demand of system reliability has received intensive attention of academics and practitioners in related areas. Among all the tasks, data-driven fault diagnosis, involving the use of data to detect and identify faults, is one of the most interesting research domains.

In previous extensively cited literature, Venkatasubramanian once proposed classical three subclasses of

* Correspondence: xiaody@mail.tsinghua.edu.cn
Department of Automation, Tsinghua University, 100084, Beijing, China

diagnostic techniques: *quantitative model-based methods, qualitative model-based methods*, and *process history based method* [3-5]. From a new perspective to further investigate Venkatasubramanian's classification, data-driven based fault diagnosis not only includes a large part of techniques in process history based method, but also some belonging to qualitative model-based methods. To view data-driven methods as an integrated type, we can re-divide fault diagnosis methods into three subclasses, namely *analytical model-based methods, qualitative knowledge-based methods*, and *data-driven based methods* (DDBM), where DDBM can be further divided into *data transform based methods* (DTBM) and *data reasoning based methods* (DRBM). Figure 1 illustrates the proposed classification. In general, DDBM are associated with the methods with insufficient information available to form mechanism model. These kinds of methods employ process data in dynamic system to perform fault detection, diagnosis, identification, and location. DTBM, to be more specifically, highlights the adoption of linear or nonlinear mathematical transforms to map original data to data in another form and the transforms are often reversible. The transformed data may be without clear physical meanings, but with more practicality. The key concept of data transform lies in two attributes: deterministic transform paradigm and realization of data compression. With this concept, the scope of DTBM is smaller and more concentrating compared to DDBM; the purpose for data utilization is more specific. DTBM also needs no in-depth knowledge about system structure as well as experience accumulation and reasoning knowledge which are necessary to DRBM. Besides, the implementation of DTBM algorithms are easily understood and realized, but the

drawback may be less robust than model based methods. Dimension transformation (often dimension reduction), filtering, decomposition and nonlinear mapping are recognized as common tools for data transform.

In Figure 1, signal processing is categorized as a data transform methodology which covers a wide range of different techniques. Typical ones are primarily filtering and multilayer signal decomposition, both requiring preset models and carefully selected parameters, like Wavelet Analysis, Hilbert-Huang Transform, *etc*. Morphological signal processing, however, gives a different viewpoint. It derives from rank-order based data sorting technique and modifies signal geometry shape to achieve filtering [6]. This feature may provide more advantages of noise reduction and detail preservation than linear tools when treating measurements in complex processes [7]. Moreover, Salembier [8] analyzed that how the performance of rank-order based filter can be adaptively optimized in terms of the filter mask and rank value. Based on the investigations above, morphological signal processing as a nonlinear data transform tool may be suitable for constructing feature extractor for pattern matching.

In our previous work (unpublished work), we developed Salembier's idea [8] to adaptively adjust flat structuring element and rank parameter for each sample rather than adopting uniform ones for all the samples in a sampled sequence. Based on this idea, we designed a signal geometry matching approach: pattern classification using one-dimensional adaptive rank-order morphological filter for fault diagnosis, named PC1DARMF approach. The proposed method belongs to DTBM with major parameters capable of being randomly chosen, which is superior to those DTBM which need
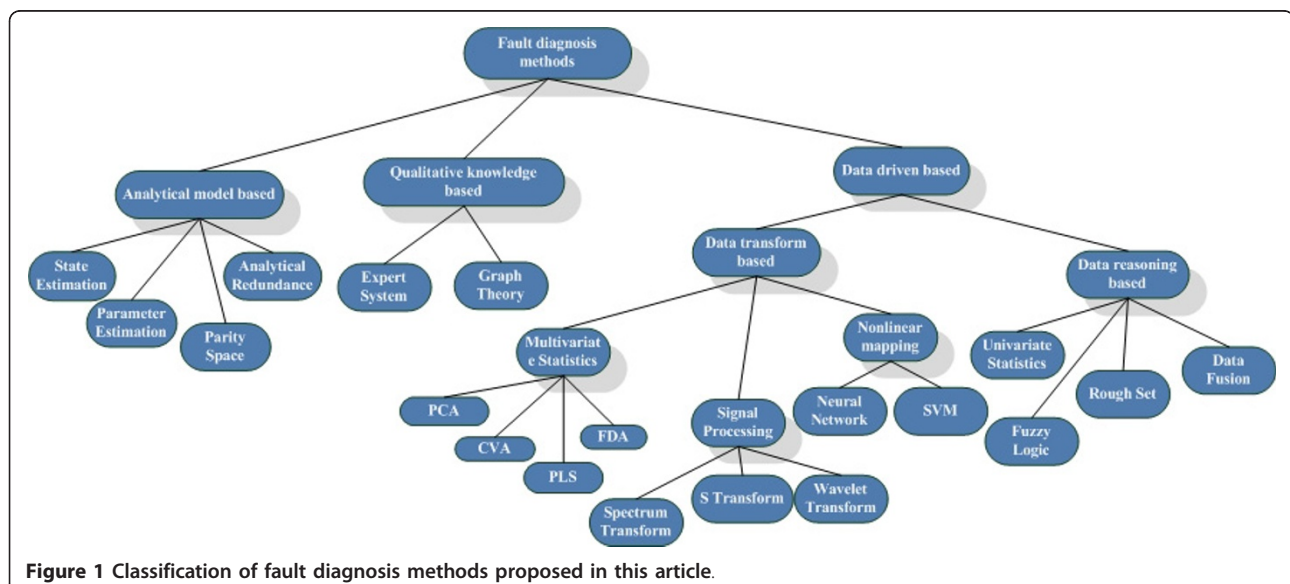

Figure 1 Classification of fault diagnosis methods proposed in this article.

predefined parameters. This article applies PC1DARMF approach to a more complex and challenging application: Tennessee Eastman process (TEP). TEP is a classic model of an industrial chemical process widely studied in literature for validating new developed control or process monitoring strategies. It is a typical large-scale process characterized by features described previously. The fact that many data-driven diagnostic methods have been performed on TEP also provides chances to evaluate their performances in comparisons with method proposed in this article.

The remainder of this article is organized as follows: Section 2 expounds the derivation of pattern classification method using adaptive rank-order morphological filter. Key implementation issues are also discussed. An example is given to build a step-by-step realization of the method, making it easier for readers to understand. Section 3 gives an essential introduction to TEP and reviews the previous TEP fault diagnosis methods. Section 4 shows the diagnosis results for different TEP simulated faults with detailed analysis. Comparisons between the proposed method and typical multivariate statistics based approaches are made to highlight the advantages and features of PC1DARMF. The last part finally presents the conclusion and discussions.

2. Signal geometry matching based on adaptive rank-order morphological filter

2.1. One-dimensional adaptive rank-order morphological filter (1DARMF)

Adaptive rank-order morphological filter is derived from a nonlinear signal processing tool referred as the rank-order based filter (ROBF). ROBF firstly reads a certain number of input values, then sorts the values in ascending order and determines the output value according to the predefined rank parameter in the sorted set. The basic definition of one-dimensional (1D) ROBF is firstly given in [9]: let x_i be discrete sampled signal defined on a 1D space Z and M be a 1D mask containing N points ($|M| = N$ and $| |$ is the set cardinality). Define j as an index belonging to the mask M and r as the normalized rank parameter of the filter ($0 \le r \le 1$). Given the rank-order operator denoted by $f_{r,M}[x_i]$, the output of ROBF y_i can be then formulated as (1):

$$y_i = f_{r,M}[x_i] = \text{Rank}_n\{x_{i-j} | j \in M\} \tag{1}$$

where elements of set X are sorted in ascending order and $\text{Rank}_n\{X\}$ denotes the nth ordered value in X (n is the nearest integer value of $(N - 1)r + 1$), x_{i-j} denote all the points which belong to the range of mask M centered on i (e.g., if $j = -3, -2, -1,0,1,2,3$, $i - j = i - 3,...,i+3$). This operation is the essentials of both median filter and morphological filter with flat structuring element [8,9].

However, its drawback is that the selections of filter mask and rank parameter heavily rely on practical experience and intuition. With understanding the feature of ROBF, its adaptive form named adaptive rank-order morphological filter was then proposed [8,9]. It is optimized as adapting filter mask and rank parameter in order to minimize a criterion such as the MAE (mean absolute error) or the MSE (mean squared error). The problem of designing adaptive rank-order morphological filter can be briefly stated as follows: assume that x_i and d_i are given as noised signal and desired signal, respectively, when ROBF $f_{r,M}$ is adopted, the aim is to find the best rank parameter r and filter mask M which minimizes a cost function C between output y_i and d_i using iterative learning. In order to expound the procedure of building 1DARMF for better understanding, how to formulate the operation of ROBF is to be introduced at the beginning.

First, in order to overcome the optimization difficulty for dealing with the discrete nature of parameters, the rank parameter r can be optimized in continuous normalized manner and let n in $\text{Rank}_n\{X\}$ be the nearest integer value of $(N - 1)r + 1$. Secondly, for filter mask M optimization problem, a search area A which is selected to be larger than the optimum mask is introduced and a continuous value $m^{(j)}$ is assigned for $\forall j \in A$. New filter mask in next iterative step is thus determined by comparing the set of continuous values associated with the current filter mask against a preset value (denoted as threshold thm_M). If the assigned value for any $j \in A$ is greater than the threshold, the location associated to that j belongs to the filter mask. With introduction of search area A and the continuous values assignments, the optimization problem of filter mask M is successfully converted from the binary values modification of the mask (belong or not belong) to continuous values $m^{(j)}$ modification.

On the basis of realizing parameters updating continuously, we proceed to find a way to establish a mathematical relationship involving filter input, output, and the parameters all together. Let us define S the sum of signs of $(x_{i-j}-y_i)$ for all j. It can be expressed by

$$S = \sum_{j \in M} \text{sgn}\ (x_{i-j} - y_i) \tag{2}$$

It is easy to find out that if $r = 0$, y_i is the minimum of $\{x_{i-j} | j \in M\}$ and S is then equal to $N - 1$; if $r = 0.5$, y_i is the median value of $\{x_{i-j} | j \in M\}$ and $S = 0$; if $r = 1$, y_i is the maximum of $\{x_{i-j} | j \in M\}$, $S = - (N - 1)$. Based on the mapping relations between S and r above, if they were assumed to be linearly related, the general expression of S with respect to r is given as

$$S = -(2r - 1)(N - 1) \tag{3}$$

In case of thm_M being set 0, we obtain if $(\text{sgn}(m^{(j)} - \text{thm\_M})+1)/2 = 1$, then $m^{(j)} > \text{thm\_M}$, which means $j \in M$ and else if $(\text{sgn}(m^{(j)}-\text{thm\_M})+1/2) = 0$, then $m^{(j)} < ^{\text{thm\_M}}$, $j \in M^c$ Notice all j is selected from A and let $(\text{sgn}(m^{(j)}-\text{thm\_M})+1/2)$ (i.e., $(\text{sgn}(m^{(j)})+1)/2$) be the weight, combing (2) and (3) gives

$$S = \sum_{j \in A} \frac{1}{2}(\text{sgn}\ (m^{(j)}) + 1)\text{sgn}\ (x_{i-j} - y_i) = -(2r-1)[\sum_{j \in A} (\text{sgn}\ (m^{(j)}) + 1)/2 - 1] \quad (4)$$

$$F(m^{(j)}, x_{i-j}, y_j, r) = \sum_{j \in A} \frac{1}{2}(\text{sgn}\ (m^{(j)}) + 1)[\text{sgn}\ (x_{i-j} - y_i) + 2r - 1] + 1 - 2r = 0 \quad (5)$$

Thus, the output of ROBF is successfully expressed by the implicit function $F(m^{(j)}, x_{i-j}, y_j, r)$. As will be stated later, this implicit function is applied to take derivatives of y_i with respect to m and r to develop iterative formulae for parameter updates.

In [8], an iterative algorithm similar to the LMS (least mean squares) algorithm was suggested to update the $m^{(j)}$ and r in the case of MSE optimization:

$$m^{(next, j)} = m^{(j)} + 2\alpha(d_i - y_i)\frac{\partial y_i}{\partial m^{(j)}} \quad \forall j \in A \quad (6)$$

$$r^{(next)} = r + 2\beta(d_i - y_i)\frac{\partial y_i}{\partial r} \quad (7)$$

Where α and β are two predefined parameters controlling the convergence rates. The derivatives of y_j with respect to $m^{(j)}$ and r are calculated through employing implicit function (5). To obtain the expression of $\frac{\partial y_i}{\partial m^{(j)}}$ and $\frac{\partial y_i}{\partial r}$, the derivative of F with respect to m_k is firstly expressed as

$$\frac{dF}{dm^{(j)}} = \frac{\partial F}{\partial m^{(j)}} + \left(\frac{\partial F}{\partial y_i}\right)\left(\frac{\partial y_i}{\partial m^{(j)}}\right) = 0 \quad (8)$$

That is

$$\frac{\partial y_i}{\partial m^{(j)}} = -\frac{\partial F/\partial m^{(j)}}{\partial F/\partial y_i} \quad (9)$$

Using (5) to take the derivative of F with respect to $m^{(j)}$ gives

$$\frac{\partial F}{\partial m^{(j)}} = \frac{\partial \text{sgn}\ (m^{(j)})}{2\partial m^{(j)}}[\text{sgn}\ (x_{i-j} - y_i) + 2r - 1]$$
$$= \delta(m^{(j)})[\text{sgn}\ (x_{i-j} - y_i) + 2r - 1] \quad (10)$$

$\frac{\partial F}{\partial y_i}$ is also calculated by using (5):

$$\frac{\partial F}{\partial y_i} = -\sum_{j \in A} (\text{sgn}\ (m^{(j)}) + 1)\delta(x_{i-j} - y_j) \quad (11)$$

In (11), the term $\delta(x_{i-j}-y_i)$ is equal to 1 only if j equals to j_0, i.e., the time shift whose corresponding x_{i-j_0} equals to output y_i. This indicates $j_0 \in M$ and $\text{sgn}(m_{j_0}) = 1$, (11) is simplified to

$$\frac{\partial F}{\partial y_i} = -2 \quad (12)$$

Combined with (10), (9) is written as

$$\frac{\partial y_i}{\partial m^{(j)}} = \frac{1}{2}\delta(m^{(j)})[\text{sgn}\ (x_{i-j} - y_i) + 2r - 1] \quad (13)$$

If $\delta(m_k)$ is replaced by $\delta'(m_k) = 1$ for $-1 \le m_k \le 1$ for simplification. Based on (13), (6) is converted to

$$m^{(next, j)} = m^{(j)} + \alpha(d_i - y_i)[\text{sgn}\ (x_{i-j} - y_i) + 2r - 1] \quad (14)$$

Similar with the deduction of (9) and (13), we also have

$$\frac{\partial y_i}{\partial r} = -\frac{\partial F/\partial r}{\partial F/\partial y_i} \quad (15)$$

$$\frac{\partial F}{\partial r} = 2\left[\frac{1}{2}\sum_{j \in A} (\text{sgn}\ (m^{(j)}) + 1) - 1\right] = 2(N - 1) \quad (16)$$

Based on (12), (16) is written as

$$\frac{\partial y_i}{\partial r} = N - 1 \quad (17)$$

Combined with (17), (7) is converted to

$$r^{(next)} = r + 2\beta(d_i - y_i)(N - 1) \quad (18)$$

where $N = |M|$ is the current length of filter mask in use.

Combining (1), (14), and (18), the parameters updating algorithm for one dimensional adaptive rank order morphological filter are given as (19), where itN denotes the current iteration and $itN + 1$ for the next. Note that the update processes of filter mask M and rank parameter r are varying according to each sample i rather than remaining the same for each sample.

$$y_i^{(itN)} = \text{Rank}_{(N_i^{(itN)}-1)r_i^{(itN)}+1} \{x_{i-j} | j \in M_i^{(itN)}\}, |M_i^{(itN)}| = N_i^{(itN)}$$

$$\begin{cases} m_i^{(itN+1),j} = m_i^{(itN),j} + \alpha(d_i - y_i^{(itN)})[\text{sgn}\,(x_{i-j} - j) - y_i^{(itN)}) + 2r_i^{(itN)} - 1], \forall j \in M_i^{(itN)} \\ M_i^{(itN+1)} = \{j | \forall j \in M_i^{(itN)}, m_i^{(itN+1),j} > \text{thm\_M}\} \\ r_i^{(itN+1)} = r_i^{(itN)} + 2\beta(d_i - y_i^{(itN)})(N_i^{(itN)} - 1) \end{cases} \quad (19)$$

To illustrate the performance of 1DARMF given by (19), an example is shown in Figure 2. In Figure 2a, it depicts three signals: noised signal x (*dash-dot* line) as input signal, desired signal d (*solid* line) as supervisory signal, and output signal y (*dotted* line) as recovered signal. $x = s + n$,

where s is the useful signal contaminated by Gaussian noise n and SNR_x (signal-to-noise ratio) is set 2. In this example, $s = \sin(t)$ and d is selected equal to s in order to recover the useful signal. Initial parameters of 1DARMF in (19) are set as follows: initial 1D filter mask $M^{(0)} = [-5,-4,-3,-2,-1,0,1,2,3,4,5]$, initial assigned value for element in the mask $m^{(0,j)} = 0.5$ ($\forall j \in M$), initial rank parameter $r^{(0)} = 0$, thm_M = 0, max iterations iterationNUM = 300, convergence rate $\alpha = 1 \times 10^{-4}$ and $\beta = 1.5 \times 10^{-3}$.
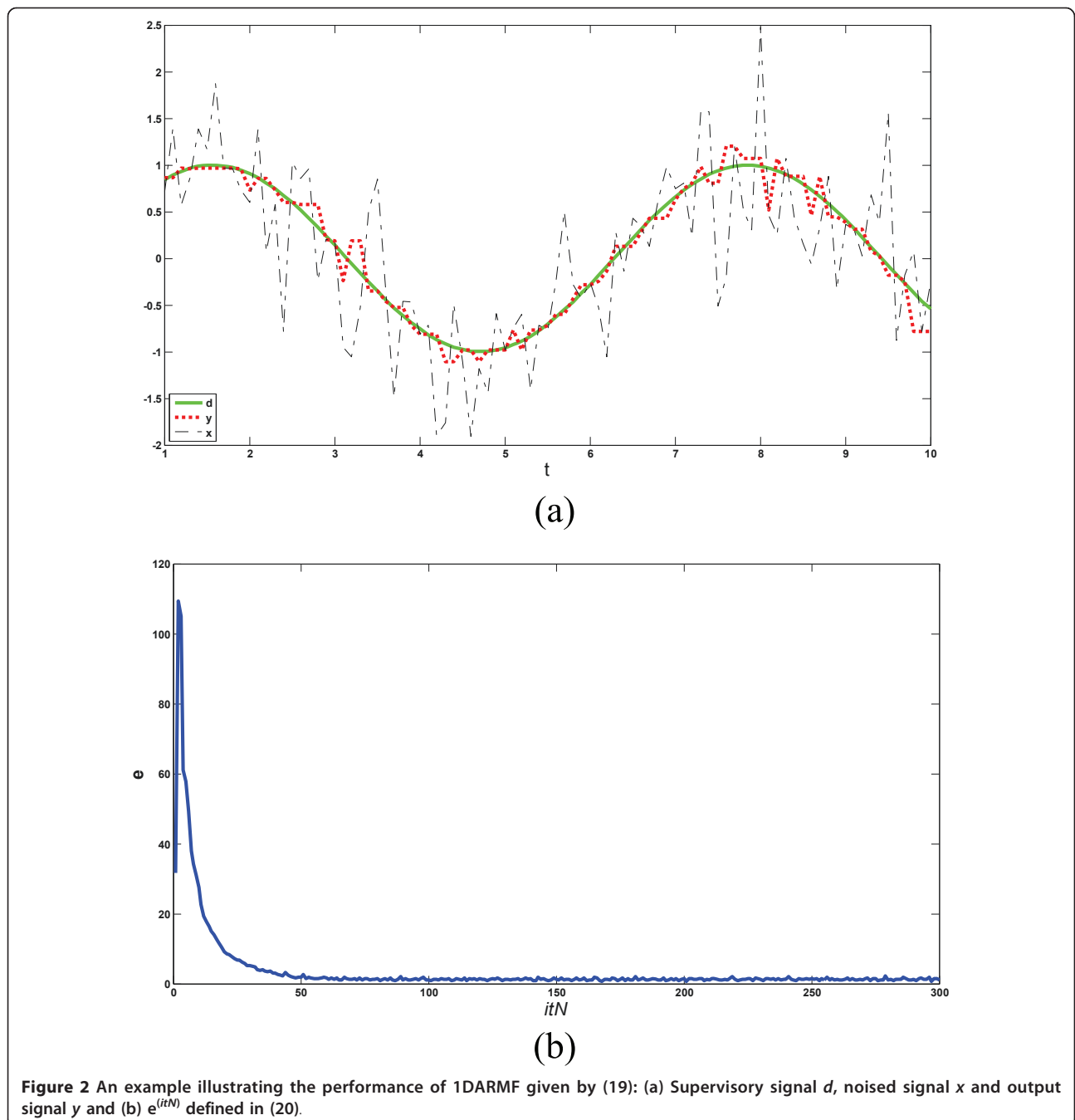

Figure 2 An example illustrating the performance of 1DARMF given by (19): (a) Supervisory signal d, noised signal x and output signal y and (b) $e^{(itN)}$ defined in (20).

If we define the sum of squared error between y and d as the evaluation of signal recovering ability, the expression is given as

$$e^{(itN)} = \sum_i |y_i^{(itN)} - d_i|^2 \qquad (20)$$

where i means the ith sample of signal and itN denotes current iteration. Figure 2b shows $e^{(itN)}$ converges to steady state and oscillates in a stable manner as itN gets increased.

2.2. Pattern classification using 1DARMF (PC1DARMF)

In Section 2.1, the general procedure to implement 1DARMF needs desired signal d as supervisory signal to train the key parameters of filter to obtain desired output. However, for a certain input x, if d is alternatively chosen, the iterative training process would finally lead to different output y. This means under supervision of inappropriate or undesirable d, the output may fail to recover useful signal from original input x. A performance comparison of 1DARMF using different supervisory signals is given to illustrate this phenomenon in Figure 3. With input x and the initial parameters being

set the same with Section 2.1, different d results in different y, as shown in Figure 3a, c, e, g, i. Figure 3b, d, f, h, j depict corresponding $e^{(itN)}$ gradually reaches stable oscillation as iterations increase. The most distinct common feature is all $e^{(itN)}$ eventually progress to a steady-state through enough iterations. This phenomenon can be theoretically guaranteed: Feuer and Weinstein [10] concluded that if the convergence rate was restrained within a upper limit, then it was the necessary and sufficient for LMS algorithm to ensure the convergence of the algorithm. Therefore, with the proper selection of α in (6) and β in (7), $e^{(itN)}$ is also expected to stably oscillate eventually. The selection rule will be later summarized in Section 2.3. This condition is the crucial prerequisite to further form our algorithm for pattern classification. In Table 1 $\min(e^{(itN)})$ are also listed to numerically compare the effect of different d on signal recovering.

Figure 3 and Table 1 indicate the most matching supervisory signal in signal geometry shape with original input x (i.e., $d = s = \sin(t)$) yields minimum value of $\min(e^{(itN)})$, showing the best signal recovering ability. Based on this property, it is expected that given an
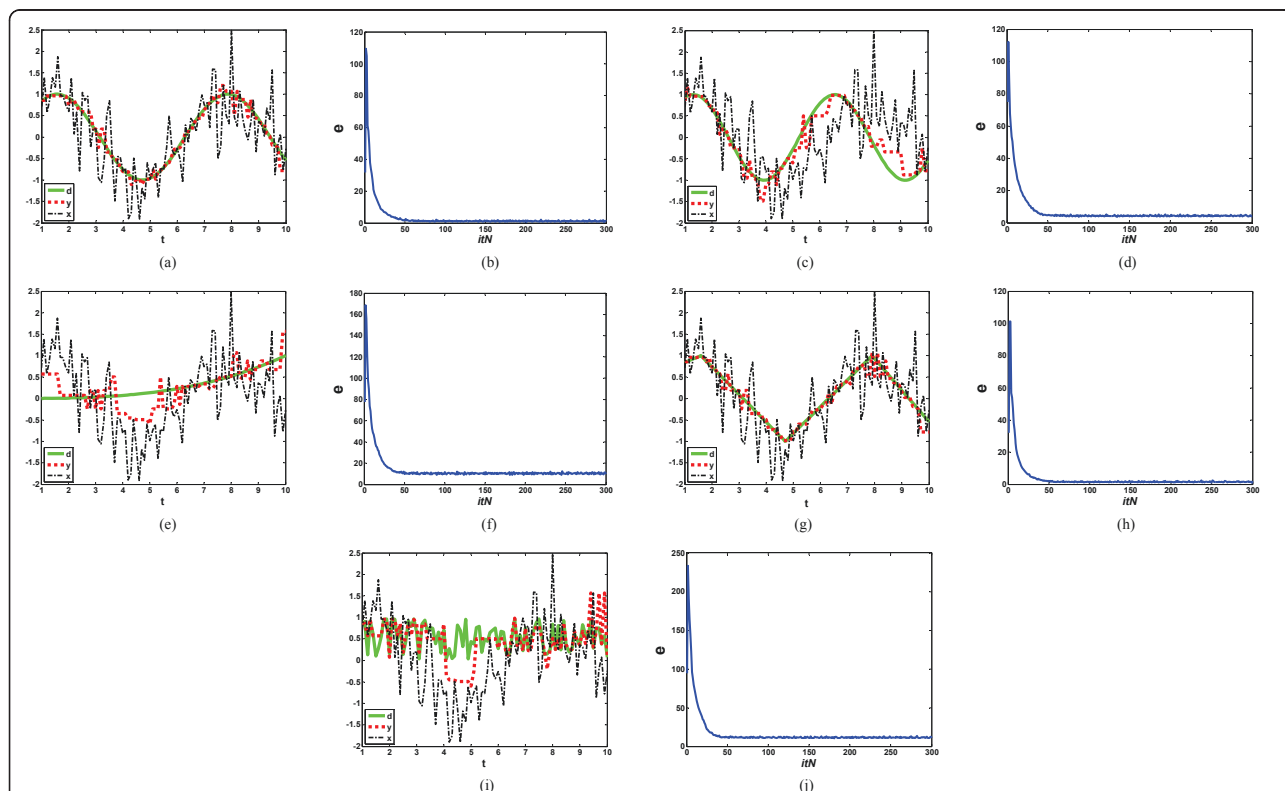
Figure 3 1DARMF performances using $s = \sin(t)$, SNR$_x$ = 2 and different supervisory signal d. Initial parameter settings: $M^{(0)}$ = [-5,-4,-3,-2,-1,0,1,2,3,4,5], $m^{(0,j)}$ = 0.5 ($\forall j \in B$), $r^{(0)}$ = 0, thm_M = 0, iterationNUM = 300; **(a)** $d = \sin(t)$, **(c)** $d = \sin(1.2t)$, **(e)** $d = c(t^3 + t^2 - 1)$ (c is a proper scaling factor which constrains range of d to be within [-1,1]), **(g)** d is triangular signal (TriWave), **(i)** d is signal generated according to uniform distribution (rand), **(b), (d), (f), (h), (j)**: correspondent $e^{(itN)}$ of its left figure.

Table 1 min(e$^{(itN)}$) gained using different supervisory signal *d* (s = sin(t))

S	D	min(e$^{(itN)}$)
sin(t)	sin(t)	0.7276
	sin(1.2t)	3.7734
	$c(t^3 + t^2 - 1)$	8.9434
	TriWave	0.9754
	Rand	10.6224

unrecognized noised signal and a certain number of reference signals (also known as signal templates) as supervisory signals, 1DARMF may be capable of achieving signal recognition and classification through finding out under which reference signal the min(e$^{(itN)}$) value reach the minimum among all reference signals provided. We thus propose the basic procedures for pattern classification using 1DARMF in Figure 4.

The procedure for pattern classification using 1DARMF can be further developed to an algorithm, named PC1DARMF algorithm. It is a supervised pattern classification approach. The fundamental of this algorithm is to realize signal geometry shape matching using 1DARMF as a tool in an iterative way. If the supervisory signals denote different types of physical meanings, for example representing different operation conditions or fault types in dynamic processes, this algorithm could achieve faults diagnosis through the signal geometry

shape matching. In general, PC1DARMF algorithm is meaningful in two levels: first, it serves for the type classification purpose and secondly a feature extractor from nonstationary signals with proper parameter settings.

2.3. Issues for implementing PC1DARMF algorithm

In Section 2.2, PC1DARMF algorithm was mainly described in a high-level structure. There are still several significant engineering principles and experience to know which are important to practical implementation. They include initial parameter settings, convergence rates selections, and iteration stopping criteria.

2.3.1. Initial parameter settings

Initial parameter settings for PC1DARMF algorithm involves initial value determination of filter mask $M^{(0)}$, assigned value $m^{(0,j)}$ for each element in filter mask, rank parameter $r^{(0)}$ and the threshold thm_M. Several reasons are supporting the random initial parameter settings. First, the only variable of filter mask in 1DARMF is its length. Based on analysis of Nikolaou and Antoniadis [11] of empirical rule for the length selection and consideration of keeping computational complexity relatively low, we propose to random chose it between 0.3 and 0.5 times of the total length of input signal. Secondly, there are no guidelines in theory for m_i and r_i initial values. They get renewal in continuous manner to optimal value during iterations, so their initial values are expected to be different chosen each time within an

Step 1: Set values of initial parameters $M^{(0)}$, $m^{(0,j)}$, $r^{(0)}$ and *thm_M*

Step2: For a input signal *x*, select a signal template d_n (*n*=1,2,3…,*Np* and *Np* is the signal templates number) as supervisory signal and apply 1DARMF until $e_n^{(itN)}$ in (20) oscillates in steady state, then calculate index FI_n=min($e_n^{(itN)}$).

Step 3: Substitute supervisory signal d_1 with d_2, d_3,……, d_{Np} respectively, repeat Step 2.

Step 4: Define *MINFI* is the minimum value of FI_n (*n*=1,2,3…,*Np*).Determine under which supervisory signal 1DARMF reaches *MINFI*. For example, if d_{n0} resulted in *MINFI*, then it indicates *x* matches signal template d_{n0} best and *x* can be classified to the corresponding group of d_{n0}.
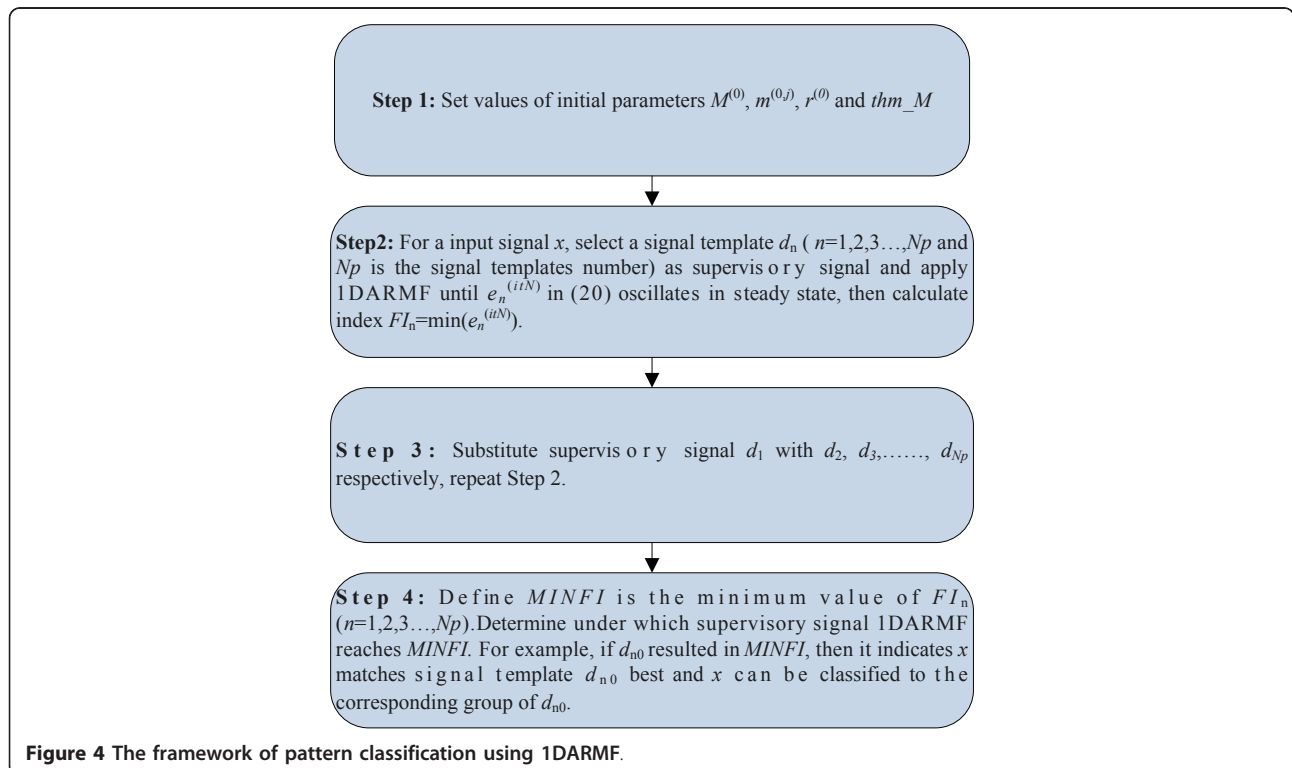
Figure 4 The framework of pattern classification using 1DARMF.

interval (e.g., [0, 1]). Thirdly, notice the derivations of (6) and (18) in Section 2.1 are all irrelevant to the value of thm_M, thm_M can be also randomly chosen within [0, 1]. Besides, the most important is that it is impossible to find optimal initial parameter settings for signals with varying nonstationary characteristics. The first goal of PC1DARMF is to measure how good two signals match each other rather than achieve optimal signal recovering, so the selection of initial parameter values would not be necessarily restrained as special ones. Based on the analysis, we use random initial parameter settings for later experiments.

2.3.2. Convergence rates selections

The selection rule of convergence rate α and β in (19) is (21), which is referenced from [10] and early mentioned in Section 2.1. As was indicated before, (21) guarantees the convergence of the LMS algorithm.

$$0 < \mu \leq \frac{1}{3tr[R]} \qquad (21)$$

where μ denotes convergence rate, R is covariance matrix of input signal, tr[R] is the trace of R. We further find empirically that if α and β is chosen as $1/3tr[R]$, output y may often cause unstable oscillation. In this article, we adopt that α and β is much smaller than $1/3tr[R]$: for example, $\alpha = 0.0001, \beta = 0.0015$.
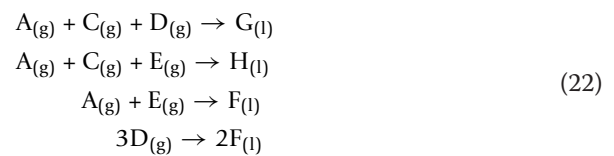
2.3.3. Iteration stop criteria

Max iteration number preset is the key factor to greatly influence the algorithm computational cost. Notice the computational complexity of PC1DARMF algorithm is $O(|\overline{N} \log \overline{N}||SL||dNUM||\text{Max}itN|)$, where $\overline{N}$ is the average length of structuring element and $O(|\overline{N} \log \overline{N}|)$ is the computational complexity of Quicksort algorithm, SL is the processed signal length, and $dNUM$ for the number of signal templates. SL and $dNUM$ are predefined and unchangeable. MaxitN is the max iterations to ensure the convergence. Salembier [8] and Figure 3 in Section 2.2 also pointed out that 1DARMF had an ability to provide fast convergence. If the PC1DARMF algorithm always set a fixed iteration numbers, it would be unnecessary and the computational cost would be tremendous. An alternative way for reducing redundant iterations is to stop the iterations when within a certain number of continuous iterations, average variation of e$^{(itN)}$ falls below a threshold if no specified information about input signal and the noise level is given.

3. Tennessee Eastman process fault diagnosis using PC1DARMF algorithm

3.1. Introduction to Tennessee Eastman process (TEP)

Tennessee Eastman process is first proposed by Downs and Vogel [12] to provide a simulated model of real industrial complex process for studying large-scale process control and monitoring methods. As is shown in Figure 5, the process consists of five major units: an exothermic two-phase reactor, a product condenser, a recycle compressor, a flash separator, and a reboiler stripper. Gaseous reactants A, C, D, E, and inert B are fed to the reactor. Component G and H are two products of TEP, while F is undesired byproduct. The reaction stoichiometry is listed as (22). All the reactions are irreversible, exothermic, and approximately first-order with respect to the reactant concentrations. The reaction rates are expressed as Arrhenius function of temperature. The reaction producing G has higher activation energy than that producing H, thus resulting in more sensitivity to temperature.

$$\begin{aligned} A_{(g)} + C_{(g)} + D_{(g)} &\rightarrow G_{(l)} \\ A_{(g)} + C_{(g)} + E_{(g)} &\rightarrow H_{(l)} \\ A_{(g)} + E_{(g)} &\rightarrow F_{(l)} \\ 3D_{(g)} &\rightarrow 2F_{(l)} \end{aligned} \qquad (22)$$

The reactor product stream is cooled through a condenser and fed to a vapor-liquid separator. The vapor exits the separator and recycles to the reactor feed through a compressor. A portion of the recycle stream is purged to prevent the inert and byproduct from accumulating. The condensed component from the separator is sent to a stripper, which is used to strip the remaining reactants. After G and H exit the base of the stripper, they are sent to a downstream process which is not included in the diagram. The inert and byproducts are finally purged as vapor from vapor-liquid separator.

The process provides 41 measured and 12 manipulated variables, denoted as XMEAS(1) to XMEAS(41) and XMV(1) to XMV(12), respectively. Their brief descriptions and units are listed in Table 2. Twenty preprogrammed faults IDV(1) to IDV(20) plus normal operation IDV(0) of TEP are given to represent different conditions of the process operation, as listed in Table 3. TEP proposed in [12] is open loop unstable and it should be operated under closed loop. Lyman and Georgakis [13] proposed a plant-wide control scheme for the process. In this article, we implement this control structure to evaluate performance of PC1DARMF algorithm on fault diagnosis for it provides the best performance for the process.

3.2. Related work for TEP fault diagnosis

Various approaches have been proposed to deal with the fault diagnosis and isolation for TEP since its introduction in 1993. Most of them are dedicated to exploit data-driven techniques because of the process complexity and data abundance. Multivariate statistics based, machine learning based, and pattern matching based methods are the most frequently adopted methodologies

Figure 5 TEP flowsheet adopting control structure proposed by [13].
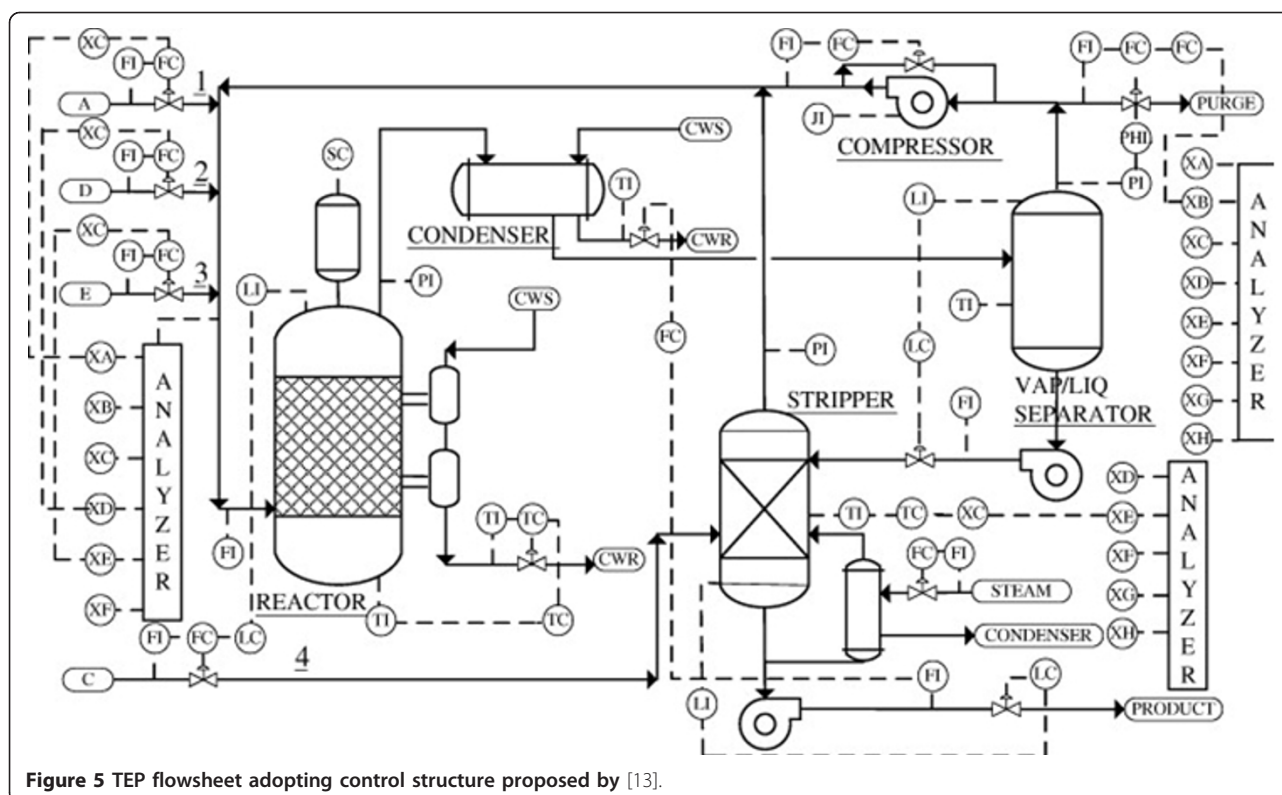
summarized in this article. Meanwhile hybrids of the three have been also studied in literature.

Raich and Cinar [14-16] are among the earliest researchers to apply multivariate statistics techniques for TEP fault diagnosis. Training data under different operation conditions are firstly utilized to design PCA (principal component analysis) models for fault detection and fault classification. Then, designed PCA models are applied to new data to calculate statistic metrics and different discriminant analysis is conducted to determine whether and which fault occurs. The method is also able to diagnosis multiple simultaneous disturbances by quantitatively measuring the similarities between models for different fault types. Russell *et al.* [17] and Chiang *et al.* [18] gives a comprehensive and detailed study of multivariate statistical process monitoring using major dimensionality reduction techniques: PCA, FDA (Fisher discriminant analysis), PLS (partial least squares), and CVA (canonical variate analysis). Additionally, some improved multivariate statistical methods outperform their conventional counterparts for TEP fault diagnosis, like dynamic PCA/FDA (DPCA/DFDA) [19], moving PCA (MPCA) [20], and modified independent component analysis (modified ICA) [21]. Application of the multivariate statistics based methods is under assumption that sample data mean and covariance are equal to their actual values [17]. This would leads to requirement

of large quantity of real data for ensuring relative accurate statistic estimations.

Machine learning based methods are also abundant in literature. It requires large amount of historical data under various fault conditions as training data to form a data mapping mechanism. Artificial neural networks (ANN) and support vector machine (SVM) are the most employed techniques applied to TEP fault diagnosis [22-25] among machine learning based methods. Eslam-loueyan [26] further proposed hierarchical artificial neural network (HANN) to diagnosis faults for TEP. Fault pattern space is first divided to subspaces using fuzzy clustering algorithm. For each subspace representing a fault pattern, a special NN is trained for fault diagnosis. Besides, Bayesian networks [27,28] and signed directed graphs (SDG) [29] are also investigated in TEP fault diagnosis problem.

Another important approach is pattern matching. The basic idea is to match the pattern against the templates stored after using feature extracting techniques. Different similarity measures are defined to quantify the matching degree. Qualitative trend analysis (QTA) is a significant pattern-matching based method. It represents signals as a set of basic shapes as major features, which distinguishes different signals in geometry shapes. Maurya *et al.* [30] used seven primitives to represent signal geometry under different fault conditions. Maurya

Table 2 Measurements and manipulated variables in TEP

Variable	Description	Units
XMEAS(1)	A feed (Stream 1)	kscmh
XMEAS(2)	D feed (Stream 2)	kg/h
XMEAS(3)	E feed (Stream 3)	kg/h
XMEAS(4)	Total feed (Stream 4)	kscm h
XMEAS(5)	Recycle flow (Stream 8)	kscm h
XMEAS(6)	Reactor feed rate (Stream 6)	kscm h
XMEAS(7)	Reactor pressure	kPa gauge
XMEAS(8)	Reactor level	%
XMEAS(9)	Reactor temperature	°C
XMEAS(10)	Purge rate (Stream 9)	kscm h
XMEAS(11)	Product sep temp	°C
XMEAS(12)	Product sep level	%
XMEAS(13)	Prod sep pressure	kPa gauge
XMEAS(14)	Prod sep underflow (Stream 10)	m³/h
XMEAS(15)	Stripper level	%
XMEAS(16)	Stripper pressure	kPa gauge
XMEAS(17)	Stripper underflow (Stream 11)	m³/h
XMEAS(18)	Stripper temperature	°C
XMEAS(19)	Stripper steam flow	kg/h
XMEAS(20)	Compressor work	kW
XMEAS(21)	Reactor cooling water outlet temp	°C
XMEAS(22)	Separator cooling water outlet temp	°C

Variable	Description	Stream
XMEAS(23)	Component A	6
XMEAS(24)	Component B	6
XMEAS(25)	Component C	6
XMEAS(26)	Component D	6
XMEAS(27)	Component E	6
XMEAS(28)	Component F	6
XMEAS(29)	Component A	9
XMEAS(30)	Component B	9
XMEAS(31)	Component C	9
XMEAS(32)	Component D	9
XMEAS(33)	Component E	9
XMEAS(34)	Component F	9
XMEAS(35)	Component G	9
XMEAS(36)	Component H	9
XMEAS(37)	Component D	11
XMEAS(38)	Component E	11
XMEAS(39)	Component F	11
XMEAS(40)	Component G	11
XMEAS(41)	Component H	11

Variable	Description	
XMV(1)	D feed flow (Stream 2)	
XMV(2)	E feed flow (Stream 3)	
XMV(3)	A feed flow (Stream 1)	
XMV(4)	Total feed flow (Stream 4)	
XMV(5)	Compressor recycle valve	
XMV(6)	Purge valve (Stream 9)	
XMV(7)	Separator pot liquid flow (Stream 10)	
XMV(8)	Stripper liquid product flow (Stream 11)	

Table 2 Measurements and manipulated variables in TEP (Continued)

XMV(9)	Stripper steam valve	
XMV(10)	Reactor cooling water flow	
XMV(11)	Condenser cooling water flow	
XMV(12)	Agitator speed	

et al. [31] also proposed an interval-halving method for trend extraction and a fuzzy matching based method for similarity estimation and inferences. Akbarya and bishnoi [32] used wavelet-based method to extract features and binary decision tree to classify them. All the above, QTA-based methods require training data, while Singhal and Seborg [33] proposed a pattern-matching-strategy requires no training data but a huge amount of historical data. The approach needs specification of snapshot dataset, which serves as a template during the historical database search. Pattern similar to snapshot data in historical database can be located by sliding a window of signals in fixed length. The drawback of this method is that it needs to accumulate historical data and, of course, cannot perform on-line process monitoring tasks. In general, pattern recognition based methods are

Table 3 Notations and descriptions of faults in TEP

Variable	Description	Type
IDV(0)	Normal operation	-
IDV(1)	A/C feed ratio, B composition constant (Stream 4)	Step
IDV(2)	B composition, A/C ratio constant (Stream 4)	Step
IDV(3)	D feed temperature (Stream 2)	Step
IDV(4)	Reactor cooling water inlet temperature	Step
IDV(5)	Condenser cooling water inlet temperature	Step
IDV(6)	A feed loss (Stream 1)	Step
IDV(7)	C header pressure loss-reduced availablity (Stream 4)	Step
IDV(8)	A, B, C feed composition (Stream 4)	Random Variation
IDV(9)	D feed temperature (Stream 2)	Random Variation
IDV(10)	C feed temperature (Stream 4)	Random Variation
IDV(11)	Reactor cooling water inlet temperature	Random Variation
IDV(12)	Condenser cooling water inlet temperature	Random Variation
IDV(13)	Reaction kinetics	Slow Drift
IDV(14)	Reactor cooling water valve	Sticking
IDV(15)	Condenser cooling water valve	Sticking
IDV(16)	Unknown	
IDV(17)	Unknown	
IDV(18)	Unknown	
IDV(19)	Unknown	
IDV(20)	Unknown	

relatively computationally demanding but the development of computer processor has been helping lessen this pressure.

Hybrid TEP fault diagnosis methods, investigated in literature commonly, employ multivariate statistical tools. For example, Lee *et al.* [34] combined SDG and PLS to demonstrate better diagnosis resolution, accuracy, and reliability than previous qualitative methods. Lu *et al.* [35] considered the limitation of PCA for differentiating faults with similar time-varying characteristics and utilized wavelet analysis to extend the feature extracting ability into time-frequency domain.

PC1DARMF algorithm in this article is a novel pattern matching method. The theoretical basis is rank-order based filter theory, which is easy to understand and implement. The computational complexity is controllable and may be relatively lower than traditional QTA-based methods. It employs the complete signal rather than some elements extracted as template. The applied strategy preserves important information, preventing

possible information loss, and distortion. It also needs no knowledge about occurrence moments of the fault. Simulation data in Section 4 would verify its effectiveness.

3.3. Diagnostic procedure of using PC1DARMF algorithm

The method proposed for TEP fault diagnosis is a supervised signal geometry shape matching approach, so constructing the signal templates as supervisory signals should be considered in the first place. The left part of Figure 6 depicts the procedure for obtaining the templates. The training data used for template construction is a matrix consisting of raw sampled signal intervals of selected measurements within different fault conditions. The normalization of the raw signal to zero mean, unit variance signal eliminates the discrepancy in the different weights given to different variables. After noise reduction by wavelet de-noising method [36], PCA model is then employed to extract PCs (principal components) since variables in TEP are highly correlated.
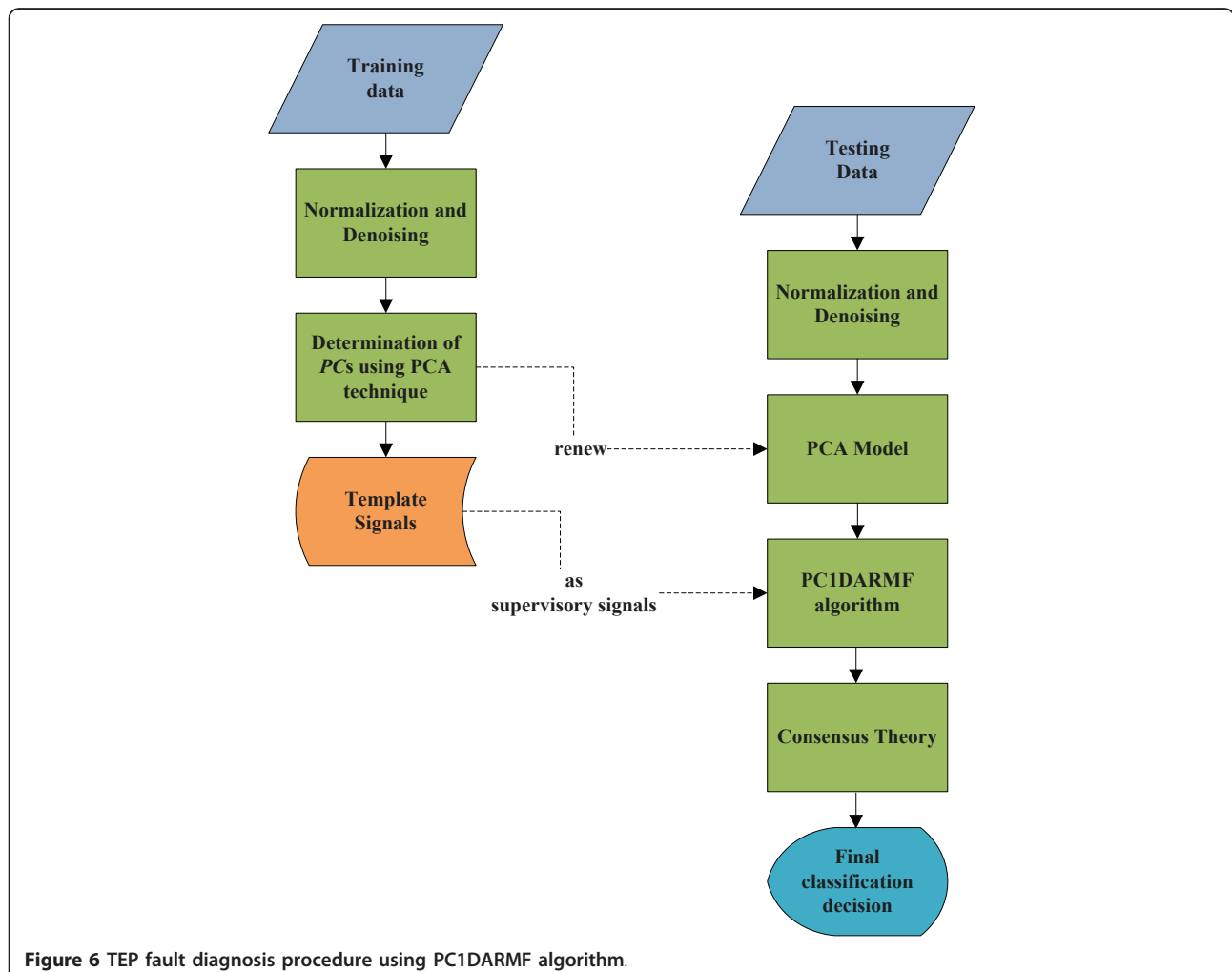

Figure 6 TEP fault diagnosis procedure using PC1DARMF algorithm.

Signals of selected PCs are defined as signal templates. This means each fault (including normal operation) is now represented as a number of signal templates. When new sampled signals as testing data are available, PCA model for training data is applied to normalized and de-noised testing signals. PC1DARMF algorithm matches unrecognized signal patterns against every template to give classification result according to each PC. Finally, multi-data fusion technique, for instance consensus theory [37], combines the classification results of selected PCs to give the final decision.

4. Simulation result analysis

4.1. Data set specification for simulation

This section describes TEP simulation data specification we adopt in this article. It mainly concerns the data constituent for the training and testing sets, data sampling interval, and sample size. The form of training and testing data is a matrix consisting of variables XMEAS (1) to XMEAS(2) and XMV(1) to XMV(11) except constant-valued XMV(12). An observation vector of TEP process is given as (23) and data collected during one simulation run of fault type i consists of k observations assembled as (24). Downs and Vogel [12] recommended the favorable time for one simulation run was between 24 and 48 hrs. In this article, we chose 24 h to keep computational cost relatively low. Russell *et al.* [17] proposed a sampling interval of 3 min to allow fast fault diagnosis. Thus, one simulation run contains 480 observations in our simulation studies, i.e., k is 480 in (24). If n_i simulation runs are implemented for fault type i, the training data matrix including n_c fault types is represented as (25).

$$x = [\text{XMEAS}(1), ..., \text{XMEAS}(41), \text{XMV}(1), ..., \text{XMV}(11)]^{\text{T}} \quad (23)$$

$$X_i = \begin{bmatrix} x_{i1}^{\text{T}} \\ ... \\ x_{ik}^{\text{T}} \end{bmatrix} \quad (24)$$

$$M_{\text{tr}} = [X_{11}, ..., X_{1n_1}, ..., X_{i1}, ..., X_{in_i}, ..., X_{n_c 1}, ..., X_{n_c n_c}]^{\text{T}} \quad (25)$$

4.2. Deterministic fault diagnosis in TEP

In this section, we first consider diagnosis of deterministic faults, i.e., IDV(1) to IDV(7) in TEP. Their brief introductions are listed in Table 3. Following the procedures described in Section 3.3, we construct the signal templates as standard patterns in the first place. The training data matrix M_{tr} of (25) is assembled. It is designed to contain process data produced in eight different simulation runs and each fault type corresponds to a simulation. In these simulations, the fault is

introduced 8 h after the simulation started. After PCA is conducted on normalized and de-noised M_{tr}, parallel analysis [38] is applied to suggest that the PCs, which correspond to the first five largest eigenvalues of training data covariance matrix, capture total variations of the data set optimally. To form standard patterns of each fault type, matrices X_0 to X_7 defined in (24) are first normalized and de-noised, respectively. After that projections of preprocessed data onto the first five PCA loading vectors are performed to obtain five new observations (scores). With the consideration of reducing the computational cost of later PC1DARMF algorithm, a resampling of the new observations reducing data points from 480 to 60 is performed. Figure 7 illustrates trends of five resampled signals for each fault type. With five in a row composing a set, these trends obtained from training data set are *signal templates*.

After signal templates were prepared, we move to build signal patterns derived from testing data. In Section 4.2, we define testing data set for each fault type, which consists of data collected from 20 simulation runs, in which 10 are generated with fault introduced in the 4th h and the 10 others introduced in the 12th h. As the same with training data set, testing data set is also needed to be normalized and de-noised. PCA model generated from training data and resampling are adopted to acquire signal patterns. The five signal patterns as a set represent current status of sampled raw data collected from the process in signal geometry manner, possessing major features for fault recognition.

After signal templates and signal patterns to be recognized are built, PC1DARMF algorithm can be performed step by step as demonstrated previously in Figure 4. The initial parameters for the algorithm are selected as follows: assigned value $m^{(0,j)}$, rank parameter $r^{(0)}$, and the threshold thm_M are subject to random choices while the initial length of filter mask $N^{(0)}$ is an integer chosen arbitrarily from 20 to 30 (0.3 to 0.5 times the trend length) for balancing tradeoff between algorithm effectiveness and low computational cost. The convergence rates α and β are pointed in Section 2.3. The algorithm stops when the variation of $e^{(itN)}$ is less than 1% within continuous 50 iterations and max iterations is set 200 to ensure algorithm convergence.

In this example, an unrecognized signal pattern of a selected PC adopts eight faulty statuses signal templates for that PC as supervisory signals and applies PC1DARMF algorithm to find its best matching signal template. If one signal pattern and its best matching signal template turn out to both derive from the same fault type, it is regarded as correct fault diagnosis, otherwise the incorrect diagnosis. Table 4 lists the correct diagnosis rates of 20 tests for each fault type using five PCs.
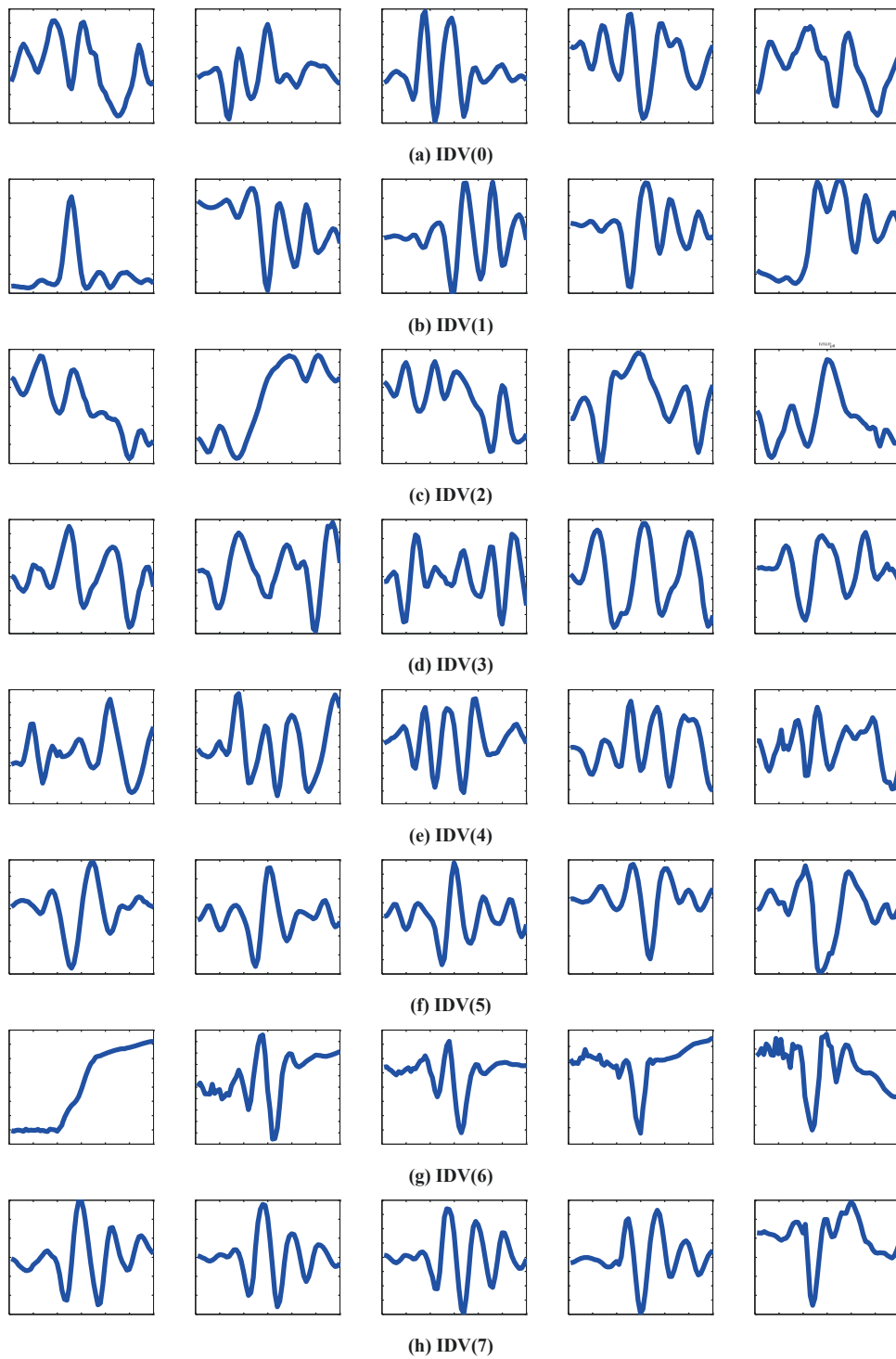
Figure 7 Set of signal templates for (a) IDV(0) to (h) IDV(7). In each set, five figures from left to right correspond to PC_1 to PC_5, respectively.

Table 4 Correct diagnosis rates of TEP deterministic faults using five PCs (20 simulations for each fault type)

PC$_i$ used:	PC$_1$	PC$_2$	PC$_3$	PC$_4$	PC$_5$
Fault type					
IDV(0)	20%	20%	5%	0.0	10%
IDV(1)	60%	100%	0.0	0.0	90%
IDV(2)	65%	100%	75%	100%	15%
IDV(3)	60%	35%	50%	30%	10%
IDV(4)	15%	10%	0.0	10%	55%
IDV(5)	75%	0.0	0.0	5%	60%
IDV(6)	100%	85%	100%	90%	30%
IDV(7)	95%	40%	0.0	5%	90%

In order to combine the fault diagnosis results given by five PCs, multisource data fusion employing consensus theory is considered here to give a more reliable final result. The linear opinion pool (LIOP) as one of the most popular approaches of consensus theory achieves results fusion by computing the weighted sum of diagnosis credibility given by each PC.

$$C(w|\text{PC}) = \sum_i \lambda_i P(w_j|\text{PC}_i) \qquad (26)$$

where $P(w_j|\text{PC}_i)$ quantifies the credibility of algorithm result of using ith PC for fault diagnosis. In other words, it reflects the frequencies of ascending orders of index FI_n (introduced in Figure 4) in overall FI values when choosing the right signal template of PC$_i$. $P(w_j|\text{PC}_i)$ can be calculated on the basis of statistics of extra training data set, which contains 10 simulation runs for each type with fault introduced time of the 8th h after the simulation started. λ_i is the weight of result given by PC$_i$ and can be determined according to data variation captured by related PC. Tables 5 and 6 list $P(w_j|\text{PC}_i)$ and weight λ_i for PC$_i$. Based on the knowledge of Tables 5 and 6 and (26), the final diagnosis results using PC1DARMF algorithm and consensus theory for IDV(0) to IDV(7) are tabulated in Table 7.

The fusion results of five PCs in Table 7 are the same with results only using signal templates of PC$_1$ in Table

Table 6 Weight λ_i of PC$_i$ for deterministic fault diagnosis

PC$_i$	λ_i
PC$_1$	60.91%
PC$_2$	15.63%
PC$_3$	11.08%
PC$_4$	7.43%
PC$_5$	4.96%

4. It requires less computational effort to only use signal templates of PC$_1$ for deterministic faults diagnosis in TEP. Table 7 suggests normal operation (IDV(0)) is correctly diagnosed with low rate. When the process is under normal operation, observations are in steady state with minor oscillation in noise level. Then, the normalization and PCA dimension reduction may result in signal templates varying randomly rather than retaining regular geometry shapes.

Table 8 compares the performances of PC1DARMF algorithm with multivariate statistics based approaches (MSBA) [17]. Both MSBA and the proposed method are on-line monitoring methods. Besides, [17] is among handful investigations which gave the detailed specifications of data in use and studied all fault types of TEP. It helps provide more comprehensive comparisons. Twenty MSBA includes PCA, DPCA, FDA.DFDA, CVA, PLS, MS (multivariate statistics) based statistic measurement (such as Hotelling T^2 or Q statistic) [17] and average values of correct diagnosis rates of 20 MSBA are listed in Table 8. It shows four out of seven faults are more easily detected by PC1DARMF algorithm rather than MSBA. IDV(3) is defined as *unobservable* from the process data in [17], which implies no observable change in mean or the variance can be detected. All MSBA performs poorly on IDV(3) diagnosis. However, PC1DARMF algorithm manages to capture variations in signal geometries and performs much better than MSBA. IDV(4) only cause the mean and standard deviation of each variable differ less than 2% between the faulty status and normal operation (IDV(0)) [17]. This phenomenon leaves signal shapes of observations almost

Table 5 Credibility of PC1DARMF algorithm for deterministic fault diagnosis using PC$_i$

Rank	PC$_1$	PC$_2$	PC$_3$	PC$_4$	PC$_5$
1	66.25%	55.00%	33.75%	41.25%	36.25%
2	16.25%	12.50%	23.75%	18.75%	25.00%
3	12.50%	21.25%	17.50%	13.75%	16.25%
4	5.00%	8.75%	8.75%	11.25%	10.00%
5	0.00	2.50%	13.75%	7.50%	6.25%
6	0.00	0.00	2.50%	6.25%	6.25%
7	0.00	0.00	0.00	1.25%	0.00
8	0.00	0.00	0.00	0.00	0.00

Table 7 IDV(0) to IDV(7) diagnosis results using PC1DARMF algorithm and consensus theory

Fault type	Correct diagnosis rate
IDV(0)	20%
IDV(1)	60%
IDV(2)	65%
IDV(3)	60%
IDV(4)	15%
IDV(5)	75%
IDV(6)	100%
IDV(7)	95%

Table 8 Comparisons of correct diagnosis rates for deterministic faults in TEP between PC1DARMF algorithm and MSBA

Fault type	PC1DARMF algorithm	MSBA [17]
IDV(1)	60%	88.76%
IDV(2)	65%	92.54%
IDV(3)	60%	12.41%
IDV(4)	15%	49.56%
IDV(5)	75%	69.44%
IDV(6)	100%	83.39%
IDV(7)	95%	75.33%
Average	67.14%	67.35%

invariable from IDV(0), causing both misclassification rates for IDV(4) and IDV(0) higher than other faults. In general, both average values for seven deterministic faults are equally well. Considering the testing data for PC1DARMF algorithm are more diverse than MSBA in [17], the proposed method fares better than the existing ones.

4.3. Stochastic fault classification in TEP

IDV(8) to IDV(12) given in Table 3 are featured by random variations in measurements when one of them occurs. In this subsection, PC1DARMF algorithm is employed to classify stochastic fault types. The training data set consists of 10 simulation runs for each fault type and fault is introduced 8 h after simulation started. The testing data set simulates faulty states with different faults occurrence time and 20 simulation runs are provided for each fault type (with fault occurrence time 4th, 2nd, 10th, 6th h per five simulations). Parallel analysis suggests that five PCs capture most variations. The initial parameter settings for algorithm follow the settings in Section 4.2. Five sets of signal templates for characterizing IDV(8) to IDV(12) are depicted in Figure 8. With the same steps in Section 4.2, Tables 9, 10, 11 and 12 list corresponding statistics for stochastic fault classification directly for saving details of derivations.

In Table 9 statistics suggest that supervision of signal template sets of every PC results in at least one 0 correct classification rate, while the fusion gives more comprehensive results. The performances of stochastic faults classification are poorer than performances of deterministic faults. One important reason was analyzed in Section 4.2 for random variations hinder formation of standard patterns. Table 13 tabulates comparisons between PC1DARMF algorithm and MSBA [17] for classifying stochastic faults in TEP. The former performs better on two out of five faults. Note that IDV(9) is also viewed as *unobservable* like IDV(3) in [17] and very difficult to be identified by MSBA. PC1DARMF algorithm here again fares better for unobservable fault. Since

random variations lead to irregular morphology of signal shape but observable quantitative variations of statistic measurements, PC1DARMF algorithm performs poorer than MSBA on average. However, the conclusion that MSBA will always give lower misclassification rates than PC1DARMF algorithm would be incorrect, because 8 out of 20 approaches studied in [17] still give lower average correct diagnosis rates compared to PC1DARMF algorithm. If more diverse training data types are provided, the classification results of the proposed method are expected to be better.

4.4. Diagnosis of all fault types in TEP

After investigating diagnostic performances of PC1DARMF algorithm for two major fault classes, respectively, we proceed to study the case when it is applied to all possible faults in TEP. The training data set consists of 10 simulation runs for each fault type and the fault is introduced the 8th h after simulation started. The testing data set simulates faulty states with different faults occurrence time and 20 simulation runs are provided for each type (with fault introduced time 4th, 2nd, 10th, 6th h per five simulations). Parallel analysis is applied to find that six PCs enough capture most variations. The parameter selection rules for PC1DARMF algorithm is the same as previous. Tables 14, 15, 16 and 17 list related statistics. Table 17 suggests that the performances of PC1ADRMF algorithm degrades when more sets of signal templates representing more fault types are provided in comparisons with Tables 7, 8, 9, 10, 11 and 12. This phenomenon implies as a supervised pattern matching strategy, PC1DARMF algorithm may require more different training data covering major features of relevant groups as much as possible to assist forming highly representative signal templates.

5. Conclusion and discussion

In this article, a supervised pattern classification method using one-dimensional adaptive rank-order morphological filter called PC1DARMF is developed to detect and recognize different faults in Tennessee Eastman process. This method generates several signals of featured geometry shapes as standard patterns on the basis of training data. With the same processing procedures as training data, testing data reflecting current operational states of TEP are transformed to signal patterns with defined specification. They are matched against standard signal patterns with employment of 1DARMF. It adaptively adjusts filter mask and rank parameter for each sample of signal rather than adopting uniform ones for all the samples. The major parameters for implementing this algorithm are capable of being randomly chosen.
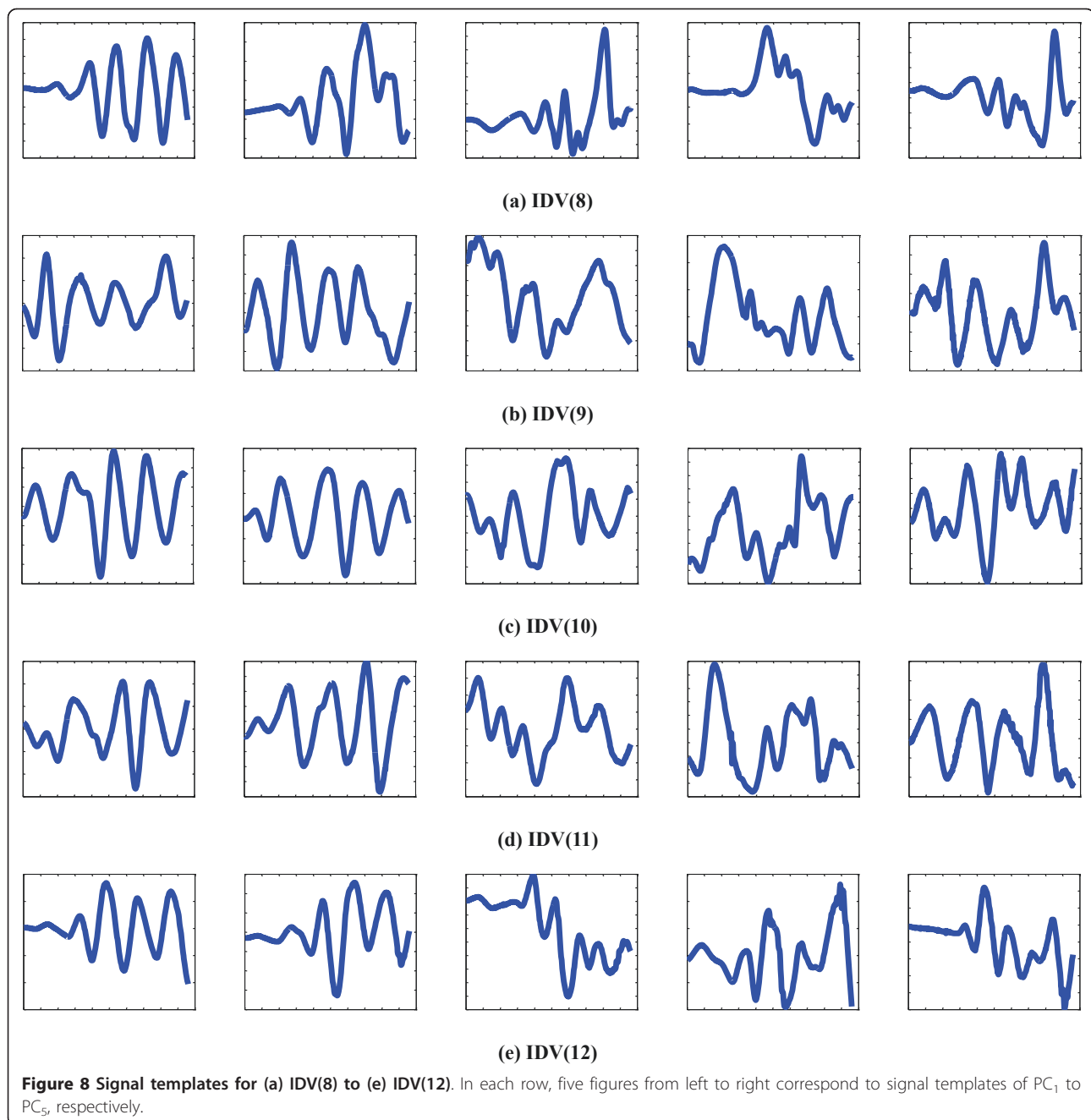
Figure 8 Signal templates for (a) IDV(8) to (e) IDV(12). In each row, five figures from left to right correspond to signal templates of PC_1 to PC_5, respectively.

Table 9 Correct classification rates of TEP stochastic faults using signal templates of five PCs (20 simulations for each fault type)

PC$_i$ used:	PC$_1$	PC$_2$	PC$_3$	PC$_4$	PC$_5$
Fault type					
IDV(8)	20%	25%	0.0	0.0	0.0
IDV(9)	40%	75%	5%	5%	45%
IDV(10)	0.0	0.0	65%	70%	5%
IDV(11)	60%	25%	30%	35%	45%
IDV(12)	20%	35%	30%	70%	0.0

Table 10 Credibility of PC1DARMF algorithm for stochastic fault classification using PC$_i$

Rank	PC$_1$	PC$_2$	PC$_3$	PC$_4$	PC$_5$
1	34.00%	32.00%	26.00%	48.00%	36.00%
2	36.00%	22.00%	40.00%	22.00%	16.00%
3	14.00%	22.00%	18.00%	12.00%	14.00%
4	10.00%	10.00%	8.00%	12.00%	20.00%
5	6.00%	14.00%	8.00%	6.00%	14.00%

Table 11 Weight λ_i of PC $_i$ for stochastic fault classification

PC$_i$	λ_i
PC$_1$	51.18%
PC$_2$	29.70%
PC$_3$	7.78%
PC$_4$	5.94%
PC$_5$	5.40%

Table 12 IDV(8) to IDV(12) classification results using PC1DARMF algorithm and consensus theory

Fault type	Correct diagnosis rate
IDV(8)	20%
IDV(9)	35%
IDV(10)	5%
IDV(11)	45%
IDV(12)	45%

TEP deterministic, stochastic, and all fault classes diagnosis are studied to verify the effectiveness of the proposed method in complex process. Consensus theory is employed for fusion of results provided by different sources The results show with only small quantity of training data provided and the testing data being much more different from training data, the performances of deterministic or stochastic faults diagnosis is better than or equally well as multivariate statistics based approaches studied in [17]. Several faults deemed as unobservable for multivariate statistic based approaches can be also recognized more easily. Deterministic faults diagnosis fares better than the stochastic faults diagnosis, since deterministic ones are apt to form similar or regular trends regardless of specified faulty conditions and noise levels while stochastic ones fail to retain basic morphologies for signal patterns. It is also noted that for some real-time applications, signal template sets provided by only one or two PCs is recommended to reduce computational cost. The future work also lies in diagnosing more diverse or multiple faults in TEP or

Table 13 Comparisons of correct diagnosis rate for stochastic faults in TEP between PC1DARMF algorithm and MSBA

Fault type	PC1DARMF algorithm	MSBA [17]
IDV(8)	20%	60.29%
IDV(9)	35%	9.52%
IDV(10)	5%	47.41%
IDV(11)	45%	36.51%
IDV(12)	45%	67.44%
Average	30%	44.23%

Table 14 Correct diagnosis rates of IDV(0) to IDV(20) using signal templates of six PCs

PC$_i$ used:	PC$_1$	PC$_2$	PC$_3$	PC$_4$	PC$_5$	PC$_6$
Fault type						
IDV(0)	5%	60%	0.0	55%	0.0	10%
IDV(1)	0.0	0.0	50%	0.0	95%	5%
IDV(2)	95%	10%	100%	100%	90%	0.0
IDV(3)	0.0	0.0	0.0	5%	15%	0.0
IDV(4)	5%	0.0	0.0	5%	0.0	0.0
IDV(5)	25%	0.0	10%	0.0	5%	45%
IDV(6)	75%	100%	100%	80%	30%	5%
IDV(7)	55%	0.0	20%	0.0	100%	85%
IDV(8)	0.0	0.0	0.0	5%	0.0	0.0
IDV(9)	0.0	20%	0.0	10%	0.0	10%
IDV(10)	0.0	0.0	0.0	0.0	5%	0.0
IDV(11)	20%	5%	0.0	5%	25%	0.0
IDV(12)	5%	0.0	0.0	0.0	0.0	55%
IDV(13)	5%	0.0	5%	5%	20%	15%
IDV(14)	5%	0.0	0.0	5%	0.0	0.0
IDV(15)	10%	5%	0.0	5%	5%	10%
IDV(16)	0.0	0.0	0.0	0.0	0.0	5%
IDV(17)	65%	50%	20%	35%	5%	10%
IDV(18)	25%	35%	65%	50%	30%	20%
IDV(19)	0.0	0.0	10%	0.0	0.0	0.0
IDV(20)	0.0	0.0	5%	0.0	10%	0.0

Table 15 Credibility of PC1DARMF algorithm for IDV(0) to IDV(20) diagnosis using PC$_i$

Rank	PC$_1$	PC$_2$	PC$_3$	PC$_4$	PC$_5$	PC$_6$
1	19.05%	17.14%	22.86%	20.48%	26.19%	20.48%
2	11.90%	6.67%	12.86%	11.43%	6.19%	8.57%
3	9.05%	5.71%	10.00%	10.00%	5.24%	4.29%
4	9.52%	7.62%	8.57%	4.76%	4.76%	4.76%
5	4.76%	5.71%	3.81%	4.76%	7.14%	7.26%
6	6.19%	7.62%	5.24%	7.14%	5.71%	5.71%
7	7.62%	8.10%	3.81%	7.14%	5.24%	6.67%
8	5.71%	5.71%	3.33%	5.71%	5.24%	4.29%
9	2.38%	7.62%	2.38%	2.86%	2.86%	2.86%
10	6.67%	5.24%	3.81%	3.81%	7.14%	4.29%
11	3.33%	5.24%	4.29%	6.19%	9.52%	1.43%
12	6.67%	5.24%	5.24%	5.24%	3.33%	4.76%
13	2.86%	1.90%	2.86%	3.81%	0.95%	2.38%
14	2.38%	5.24%	3.81%	0.48%	2.38%	3.81%
15	0.48%	0.48%	2.38%	0.95%	0.95%	2.86%
16	0.48%	1.43%	1.90%	1.90%	3.33%	3.81%
17	0.48%	1.43%	0.48%	1.90%	0.95%	5.24%
18	0.0	0.95%	0.48%	1.43%	1.90%	3.81%
19	0.0	0.0	1.43%	0.0	0.95%	1.43%
20	0.0	0.48%	0.48%	0.0	0.0	0.95%
21	0.48%	0.48%	0.0	0.0	0.0	0.0

Table 17 IDV(0) to IDV(20) diagnosis results using PC1DARMF algorithm

Fault type	Correct diagnosis rate
IDV(0)	15%
IDV(1)	30%
IDV(2)	95%
IDV(3)	0.0
IDV(4)	5%
IDV(5)	25%
IDV(6)	100%
IDV(7)	65%
IDV(8)	0.0
IDV(9)	5%
IDV(10)	0.0
IDV(11)	15%
IDV(12)	0.0
IDV(13)	5%
IDV(14)	5%
IDV(15)	5%
IDV(16)	0.0
IDV(17)	85%
IDV(18)	30%
IDV(19)	0.0
IDV(20)	0.0

other complex processes to improve the proposed methods. Besides, Ku *et al.* [19] pointed DPCA model was expected to perform better than regular PCA on the TEP problem. DPCA could be introduced to extract more information in data set.

Despite some promising results of the proposed method, this article presents only a preliminary implementation in complex process, which sheds light on the novel idea of PC1DARMF. The key concept of this idea is to achieve pattern matching between standard pattern and unrecognized pattern by 1DARMF. The specification of pattern defined is not only limited to time-domain signal geometry shapes or even not signal geometry shapes. Frequency spectrums, power spectrum, vector bases, and feature coefficients, *etc.* can be assembled to form user-defined pattern as well. The new developed pattern should be capable of not only capturing different characteristics for each group but

Table 16 Weight λ_i of PC_i for IDV(0) to IDV(20) diagnosis

PC_i	λ_i
PC_1	41.07%
PC_2	27.48%
PC_3	13.03%
PC_4	9.60%
PC_5	4.60%
PC_6	4.23%

also maintaining a relatively steady form without too much unexpected random variations. Using the same scheme may address the problems such as unstable patterns introduced by random variations.

List of abbreviations
ANN: artificial neural networks; CVA: canonical variate analysis; DDBM: data-driven based methods; DRBM: data-reasoning based methods; DTBM: data-transform based methods; D/PCA: dynamic/principal component analysis; D/FDA: dynamic/fisher discriminant analysis; 1DARMF: one dimensional adaptive rank-order morphological filter; HANN: hierarchical artificial neural network; ICA: independent component analysis; LMS: least mean squares; MAE: mean absolute error; MPCA: moving PCA; MSBA: multivariate statistics based approaches; MSE: mean squared error; PC1DARMF: pattern classification using one dimensional adaptive rank-order morphological filter; PI: proportional-integral; PID: proportional-integral-derivative; PLS: partial least squares; QTA: qualitative trend analysis; ROBF: rank-order based filter; SDG: signed directed graphs; SNR: signal-to noise ratio; SVM: support vector machine; TEP: Tennessee Eastman process.

References
1. M Iri, K Aoki, E O'Shima, H Matsuyama, An algorithm for diagnosis of system failures in the chemical process. Comput Chem Eng. 3(1-4), 489–493 (1979). doi:10.1016/0098-1354(79)80079-4
2. MA Paulonis, JW Cox, A practical approach for large-scale controller performance assessment, diagnosis, and improvement. J Process Control. 13(2), 155–168 (2003). doi:10.1016/S0959-1524(02)00018-5
3. V Venkatasubramanian, R Rengaswamy, K Yin, SN Kavuri, A review of process fault detection and diagnosis part I: quantitative model-based methods. Comput Chem Eng. 27(3), 293–311 (2003). doi:10.1016/S0098-1354(02)00160-6
4. V Venkatasubramanian, R Rengaswamy, SN Kavuri, A review of process fault detection and diagnosis part II: qualitative models and search strategies. Comput Chem Eng. 27(3), 313–326 (2003). doi:10.1016/S0098-1354(02)00161-8
5. V Venkatasubramanian, R Rengaswamy, SN Kavuri, K Yin, A review of process fault detection and diagnosis part III: process history based methods. Comput Chem Eng. 27(3), 327–346 (2003). doi:10.1016/S0098-1354(02)00162-X
6. R Stevenson, G Arce, Morphological filters Statistics and further syntactic properties. IEEE Trans Circuits Syst. 34(11), 1292–1305 (1987). doi:10.1109/TCS.1987.1086067
7. I Pitas, AN Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications* (Kluwer Academic Publishers, Boston, 1990), p. 1
8. P Salembier, Adaptive rank order based filters. Signal Process. 27(1), 1–25 (1992). doi:10.1016/0165-1684(92)90108-9
9. P Salembier, M Kunt, Multiresolution decomposition and adaptive filtering with rank order based filters–application to defect detection, in *Proceedings of IEEE International Conference Acoustics, Speech and Signal Process*, (Toronto, Canada) 2389–2392 (1991)
10. A Feuer, E Weinstein, Convergence analysis of LMS filters with uncorrelated Gaussian data. IEEE Trans Acoust Speech Signal Process. 33(1), 222–230 (1985). doi:10.1109/TASSP.1985.1164493

11. NG Nikolaou, IA Antoniadis, Application of morphological operators as envelope extractors for impulsive-type periodic signals. Mech Syst Signal Process. **17**(6), 1147–1162 (2003). doi:10.1006/mssp.2002.1576

12. JJ Downs, EF Vogel, A plant-wide industrial process control problem. Comput Chem Eng. **17**(3), 245–255 (1993)

13. PR Lyman, C Georgakis, Plant-wide control of the Tennessee Eastman problem. Comput Chem Eng. **19**(3), 321–331 (1995). doi:10.1016/0098-1354 (94)00057-U

14. A Raich, A Cinar, Multivariate statistical methods for monitoring continuous processes: assessment of discrimination power of disturbance models and diagnosis of multiple disturbances. Chemomet Intel Lab Syst. **30**(1), 37–48 (1995). doi:10.1016/0169-7439(95)00035-6

15. A Raich, A Cinar, Statistical process monitoring and disturbance diagnosis in multivariate continuous processes. AIChE J. **42**(4), 995–1009 (1996). doi:10.1002/aic.690420412

16. A Raich, A Cinar, Diagnosis of process disturbances by statistical distance and angle measures. Comput Chem Eng. **21**(6), 661–673 (1997). doi:10.1016/S0098-1354(96)00299-2

17. EL Russell, LH Chiang, RD Braatz, *Data-driven methods for fault detection and diagnosis in chemical processes* (Springer, New York, 2000) pp. 13–162

18. LH Chiang, EL Russell, RD Braatz, Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis. Chemomet Intel Lab Syst. **50**(2), 243–252 (2000). doi:10.1016/S0169-7439(99)00061-1

19. W Ku, RH Storer, C Georgakis, Disturbance detection and isolation by dynamic principal component analysis. Chemomet Intel Lab Syst. **30**(1), 179–196 (1995). doi:10.1016/0169-7439(95)00076-3

20. M Kano, K Nagao, S Hasebe, I Hashimoto, H Ohno, R Strauss, BR Bakshi, Comparison of statistical process monitoring methods: application to the Tennessee Eastman challenge problem. Comput Chem Eng. **26**(2), 161–174 (2002). doi:10.1016/S0098-1354(01)00738-4

21. J-M Lee, S Joe Qin, Fault detection and diagnosis based on modified independent component analysis. AIChE. **52**(10), 3501–3514 (2006). doi:10.1002/aic.10978

22. J Chen, C-M Liao, Dynamic process fault monitoring based on neural network and PCA. J Process. **12**(2), 277–289 (2002). doi:10.1016/S0959-1524 (01)00027-0

23. MN Nashalji, MA Shoorehdeli, M Teshnehlab, Fault detection of the Tennessee Eastman process using improved PCA and neural classifier. Soft computing in industrial applications **75**, 41–50 (2010). doi:10.1007/978-3-642-11282-9_5

24. LH Chiang, ME Kotanchek, AK Kordon, Fault diagnosis based on Fisher discriminant analysis and support vector machines. Comput Chem Eng. **28**(8), 1389–1401 (2004)

25. A Kulkarni, VK Jayaraman, BD Kulkarni, Knowledge incorporated support vector machines to detect faults in Tennessee Eastman process. Comput Chem Eng. **29**(10), 2128–2133 (2005)

26. R Eslamloueyan, Designing a hierarchical neural network based on fuzzy clustering for fault diagnosis of the Tennessee-Eastman process. Appl Soft Comput. **11**(1), 1407–1415 (2011). doi:10.1016/j.asoc.2010.04.012

27. S Verron, T Tiplica, A Kobi, Distance rejection in a bayesian network for fault diagnosis of industrial systems, in *Proceedings of 16th Mediterranean Conference on Control and Automation*, Ajaccio, France 615–620 (2008)

28. S Verron, T Tiplica, A Kobi, Fault diagnosis with bayesian networks: application to the Tennessee Eastman Process, in *Proceedings of 2006 IEEE International Conference on Industrial Technology*, 98–103 (2006)

29. MR Maurya, R Rengaswamy, V Venkatasubramanian, Application of signed digraphs-based analysis for fault diagnosis of chemical process flowsheets. Eng Appl Artif Intell. **17**(5), 501–518 (2004). doi:10.1016/j.engappai.2004.03.007

30. MR Maurya, R Rengaswamy, V Venkatasubramanian, Fault diagnosis by qualitative trend analysis of the principal components. Chem Eng Res Des. **83**(9), 1122–1132 (2005). doi:10.1205/cherd.04280

31. MR Maurya, R Rengaswamy, V Venkatasubramanian, Fault diagnosis using dynamic trend analysis: a review and recent developments. Eng Appl Artif Intell. **20**(2), 133–146 (2007). doi:10.1016/j.engappai.2006.06.020

32. F Akbaryan, PR Bishnoi, Fault diagnosis of multivariate systems using pattern recognition and multisensor data analysis technique. Comput Chem Eng. **25**(9-10), 1313–1339 (2001). doi:10.1016/S0098-1354(01)00701-3

33. A Singhal, DE Seborg, Evaluation of a pattern matching method for the Tennessee Eastman challenge process. J Process Control. **16**(6), 601–613 (2006). doi:10.1016/j.jprocont.2005.10.005

34. G Lee, C Han, ES Yoon, Multiple-fault diagnosis of the Tennessee Eastman Process based on system decomposition and dynamic PLS. Ind Eng Chem Res. **43**, 8037–8048 (2004). doi:10.1021/ie049624u

35. N Lu, F Wang, F Gao, Combination method of principal component and wavelet analysis for multivariate process monitoring and fault diagnosis. Ind Eng Chem Res. **42**, , 4198–4207 (2003)

36. DL Donoho, De-nosing by soft-thresholding. IEEE Trans Inf Theory **41**(3), 613–627 (1995). doi:10.1109/18.382009

37. JA Benediktsson, PH Swain, Consensus theoretic classification methods. IEEE Trans Syst Man Cybern. **22**(4), 688–704 (1995)

38. J Edward Jackson, *A User's Guide to Principal Components* (Wiley, New York, 2003), pp. 46–47