

# 基于动态特性描述的变量加权型分散式故障检测方法

钟凯<sup>1</sup> 韩敏<sup>1</sup> 韩冰<sup>2</sup>

**摘要** 现代工业生产过程往往具有复杂的动态特性: 不同测量变量间会存在不同的时序相关性, 且变量间的相互影响会反映在不同的采样时刻上. 现有的动态过程监测模型往往不能充分挖掘变量间的动态特性, 其故障检测效果也有待进一步提高. 在此背景下, 本文提出一种基于动态特性描述的变量加权型分散式故障检测方法. 利用最大相关最小冗余 (Minimal redundancy maximal relevance, mRMR) 算法更准确地描述动态过程变量间的相关性关系, 并利用该相关性的值对原始增广矩阵进行加权处理, 且不同延迟变量对当前测量值的影响大小就通过权值来体现, 因此能更加全面地刻画该测量值的动态特性. 最后建立一种融合 mRMR 算法, 贝叶斯推理以及动态主成分分析 (Dynamic principal component analysis, DPCA) 模型的新的分布式建模策略, 提高了模型的容错能力和泛化能力, 取得了更好的故障检测结果.

**关键词** 主成分分析, 故障检测, 最大相关最小冗余, 加权型分散式方法

**引用格式** 钟凯, 韩敏, 韩冰. 基于动态特性描述的变量加权型分散式故障检测方法. 自动化学报, 2021, 47(9): 2205–2213

**DOI** 10.16383/j.aas.c180276

## Dynamic Feature Characterization Based Variable-weighted Decentralized Method for Fault Detection

ZHONG Kai<sup>1</sup> HAN Min<sup>1</sup> HAN Bing<sup>2</sup>

**Abstract** Modern industrial process is often accompanied by complicated dynamic behaviours: different measured variables have different serial correlations, and the interactions among these variables are reflected in different sampling instants. The existing dynamic process monitoring models can not fully excavate the dynamic characteristics among variables, then the corresponding fault detection results need to be further improved. Against this background, this paper proposes the dynamic feature characterization based variable-weighted decentralized method for fault detection, which uses the minimal redundancy maximal relevance (mRMR) algorithm to get more accurate description of the relationship among different measured variables in dynamic process, then the augmented matrix can be weighted according to the mRMR values, and the influences of different delay variables on current measured one are reflected by the weights. Hence, the dynamic characteristics of the measured variable can be described more comprehensively. Finally, a new distributed modeling strategy combining the mRMR algorithm, Bayesian inference and dynamic principal component analysis (DPCA), which improves fault tolerance and generalization ability of the model, thus obtaining better fault detection results.

**Key words** Principal component analysis, fault detection, minimal redundancy maximal relevance (mRMR), weighted decentralized method

**Citation** Zhong Kai, Han Min, Han Bing. Dynamic feature characterization based variable-weighted decentralized method for fault detection. *Acta Automatica Sinica*, 2021, 47(9): 2205–2213

随着现代工业水平的迅猛提高, 其生产规模和生产过程的复杂性也大大增加, 为了保障工业安全

和产品质量, 过程监控和故障诊断技术也显得愈发重要<sup>[1-2]</sup>. 传统的基于模型的过程监测方法依赖于精确的数学模型或充分的先验知识, 极大地限制了其在实际中的应用. 而基于数据驱动的方法无需建立精确的过程机理模型, 直接对过程运行数据进行分析和处理, 据此建立描述过程运行状态的统计模型, 因此很适合于监测难以建模的复杂工业过程, 主要包括, 适合于故障分类和诊断的费舍尔判别 (Fisher discriminant analysis, FDA)<sup>[3]</sup>, 主成分分析 (Principal component analysis, PCA)<sup>[4]</sup>, 利用质量变量引导过程变量样本空间分解的偏最小二乘 (Partial

收稿日期 2018-05-03 录用日期 2018-12-12  
Manuscript received May 3, 2018; accepted December 12, 2018  
国家自然科学基金 (61773087) 资助  
Supported by National Natural Science Foundation of China (61773087)  
本文责任编辑 邓方  
Recommended by Associate Editor DENG Fang  
1. 大连理工大学电子信息与电气工程学部 大连 116023 2. 上海船舶运输科学研究所航运技术与安全国家重点实验室 上海 200135  
1. Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116023 2. State Key Laboratory of Navigation and Safety Technology, Shanghai Ship and Shipping Research Institute, Shanghai 200135

least squares, PLS)<sup>[5]</sup> 以及适用于非高斯过程监测的独立主元分析 (Independent component analysis, ICA)<sup>[6]</sup>. 特别地, PCA 作为一种最主要的数据驱动方法, 能够有效地进行数据降维和潜在特征的提取, 并且在上述方法中具有最低的计算复杂度, 因此受到工业界和学术界的广泛关注, 并在故障检测和诊断领域取得了丰富的研究成果<sup>[7-8]</sup>.

实际工业过程采集的过程数据一般都在时间上序列相关, 而基于传统的 PCA 方法的过程监控并没有考虑数据间的此种关联, 以至于主成分甚至残差都具有未被建模的动态特性, 因此并不适用于动态过程的建模. 文献 [9] 首先使用延时测量值扩充分析数据阵, 然后对该增广数据阵实施 PCA 分解以建立数据间的动态关联, 进而提出了动态 PCA (Dynamic PCA, DPCA) 的主体框架. 此后, DPCA 在理论和应用上仍有后继的发展<sup>[10-11]</sup>. 最近, 文献 [12] 利用自回归模型提取数据间动态潜隐成分 (dynamic latent variables, DLV), 而 PCA 模型用于建立残差的静态关系, 因此静态关系和动态关系都得到了很好的表征, 从而提高了动态过程的故障检测效果.

然而, 以上所涉及的方法都是单个集中的模型, 不适用于目前大规模的工业过程故障检测, 而分散式建模策略由于其具有较低的计算复杂度, 较好的模型灵活性和较强的容错能力等优点, 已成为监控大规模工业过程的有力手段<sup>[13-14]</sup>. 更进一步, 文献 [15] 利用动态特性选择准则并结合分布式建模的优势, 成功地解决了大规模动态的故障检测问题. 但其在选择子模型变量时所依赖的动态特性准则是一种单纯的线性关系指标, 无法评估变量之间的非线性关联, 因此并没有取得最优的故障检测效果. 文献 [16] 又提出一种基于互信息 (Mutual information, MI) 的分散式动态 PCA 故障检测方法 (MI-DPCA), MI 准则同时考虑了变量间的线性和非线性关系, 较好地处理了数据的动态性问题. 然而在分散式子块中基于 MI 准则选择的变量之间往往存在较大的冗余, 从而导致欠佳的故障检测结果. 文献 [17] 提出了一种基于最大相关最小冗余 (Minimal redundancy maximal relevance, mRMR) 变量选择的 PCA 故障检测, 解决了由 MI 准则带来的变量间冗余问题, 取得了较好的故障检测效果. 但该模型采取简单地剔除 mRMR 值超过设定阈值的变量来实现变量选择以建立相应的子块, 因此直接忽略了对当前变量影响较小的测量变量, 并且也不能定量描述保留下的不同测量变量对当前变量的影响程度, 此外该模型也无法解决动态过程的监控问题.

基于以上分散式建模的优势以及描述变量间动态关系方法的不足的讨论, 并受到文献 [18] 的启发,

本文提出了一种基于动态特性描述的变量加权型分散式故障检测方法 (Minimal redundancy maximal relevance — weighted dynamic principal component analysis mRMR-WDPCA). 全文主要创新点如下:

1) 与传统的变量选择方法<sup>[15-16]</sup> 相比, mRMR 算法考虑了动态过程中变量间的线性和非线性关联的同时也显著地减少变量间存在的冗余, 从而能更准确地描述变量间的相关性关系.

2) 根据变量间 mRMR 值的大小对动态增广矩阵中所有测量值赋予不同的权值, 避免遗失含有有用信息的变量, 且不同延迟变量对当前测量值的影响大小就通过权值来体现, 因此能更加全面地刻画该测量值的动态特性.

3) 提出一种融合 mRMR 算法, 贝叶斯推理以及 DPCA 模型的加权型分布式建模策略, 提高了模型的容错能力和泛化能力, 取得了更好的故障检测结果. 为全流程动态过程监控方案的设计提供了新的思路.

## 1 基本方法理论

### 1.1 PCA 和 DPCA 模型

PCA 是一种典型的线性数据降维手段, 它通过最大方差准则将原始的变量转化为新的互不关联的变量, 其中具有较大方差的新的变量称为主成分, 而方差较小的变量则是原始数据在残差空间的投影. 因此, PCA 将原始数据分解成主子空间和残差子空间. 具体地, 对于一含有  $n$  个样本和  $m$  个变量的原始数据矩阵  $X \in \mathbf{R}^{n \times m}$ , PCA 分解具有如下的数学表达形式:

$$X = \hat{X} + \tilde{X} = \sum_{i=1}^k t_i p_i^T + \tilde{X} \quad (1)$$

其中,  $t_i$  是第  $i$  个主成分,  $p_i$  是相应的载荷向量,  $k$  是保留的主元个数, 且  $\hat{X}$  是通过主元重构的矩阵,  $\tilde{X}$  是残差矩阵.

相比于常规的静态 PCA 模型, DPCA 首先对当前的数据引入前  $d$  个延时测量值, 将原始数据矩阵  $X$  转化为如下的动态增广矩阵  $X_a$ :

$$X_a = \begin{bmatrix} x_{1+d}^T & x_d^T & \cdots & x_1^T \\ x_{2+d}^T & x_{1+d}^T & \cdots & x_2^T \\ \vdots & \vdots & \ddots & \vdots \\ x_n^T & x_{n-1}^T & \cdots & x_{n-d}^T \end{bmatrix} \in \mathbf{R}^{(n-d) \times m(d+1)} \quad (2)$$

其中,  $x_i$  ( $i = 1, 2, \dots, n$ ) 表示第  $i$  个采样时刻的样

本数据. 再将常规的 PCA 方法实施于  $X_a$  即可建立相应地 DPCA 模型.

## 1.2 MI 和 mRMR

MI 是信息论中评估测量变量间相关性的测度指标. 设两个随机变量  $x_1$  和  $x_2$  的边缘概率分别为  $p(x_1)$ ,  $p(x_2)$ , 联合分布为  $p(x_1, x_2)$ , 则两者间的互信息  $I(x_1, x_2)$  为<sup>[19]</sup>

$$I(x_1, x_2) = \iint_{x_1, x_2} p(x_1, x_2) \lg \frac{p(x_1, x_2)}{p(x_1)p(x_2)} dx_1 dx_2 \quad (3)$$

文献 [20] 在互信息的基础上进一步拓展, 提出了 mRMR 算法, 其基本思想就是考虑变量间的相关性和冗余性. 集合  $S$  中变量和变量  $c$  的最大相关性定义为

$$\begin{aligned} \max D(S, c) \\ \text{s. t. } D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i, c) \end{aligned} \quad (4)$$

其中,  $|S|$  为集合  $S$  中变量的个数. 由最大相关性可知选择的变量间可能存在较大的冗余, 而最小冗余准则可由式 (5) 定义为

$$\begin{aligned} \min R(S) \\ \text{s. t. } R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i, x_j) \end{aligned} \quad (5)$$

给出下面的简单算子  $\Phi(D, R)$ , 来同时优化  $D$  和  $R$

$$\begin{aligned} \max \Phi(D, R) \\ \text{s. t. } \Phi = D - R \end{aligned} \quad (6)$$

而在实际中, 下面的增量方法常用来搜索满足式 (6) 的最优特征

$$\max_{x_j \in X - S_{m-1}} \left[ I(x_j, c) - \frac{1}{m-1} \sum_{x_i \in S_{m-1}} I(x_j, x_i) \right] \quad (7)$$

## 2 mRMR-WDPCA 的故障监测方法

### 2.1 基于 mRMR 动态特性描述

由于大规模动态过程变量间存在复杂的相关性. 为了更好地描述每个测量变量的动态特性, 提出如图 1 所示的动态特性描述方法. 具体地,  $x_t^i$  ( $i = 1, 2, \dots, m$ ) 表示  $X_t$  中的第  $i$  个测量变量, 计算其与  $X_a$  中所有变量  $x_l^j$  ( $l = t, t-1, \dots, t-d$ ,  $j = 1, 2, \dots, m$ ) 之间的 mRMR 值, 即

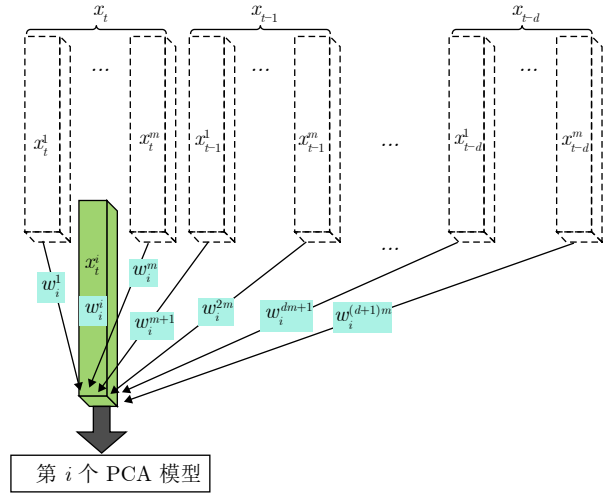


图 1 基于 mRMR 的动态特性描述

Fig. 1 mRMR-based dynamic feature characterization

$$\Phi_{i,j} = \Phi(x_t^i, x_l^j) \quad (8)$$

对矩阵  $\Phi_{i,j}$  中各元素处理得到增广矩阵  $X_a$  中各变量的权值  $w_i^j = \Phi_{i,j} / \Phi_{i,i}$ . 再对矩阵  $X_a$  进行加权处理得到  $X_i$ :

$$X_i = X_a \times \text{diag}\{w_i^1, w_i^2, \dots, w_i^{(d+1)m}\} \quad (9)$$

其中,  $[w_i^1, w_i^2, \dots, w_i^{(d+1)m}]$  为权值向量  $\mathbf{W}_i$ , 最后对求得的  $m$  个不同的加权矩阵  $(X_1, X_2, \dots, X_m)$  分别建立基于 PCA 的故障检测模型, 即

$$X_i = T_i P_i^T + E_i \quad (10)$$

由上述描述可知, 与常规的 DPCA 对所有的  $(d+1)m$  个变量赋相同的权值, 亦或是与其他直接剔除某些相关性较弱的变量选择方法相比, 式 (9) 加权了所有的变量, 在避免信息损失的同时, 还通过较大的权值强调了相关程度较大的变量带来的“积极作用”, 而相对小的权值也抑制相关程度较小的延迟变量所带来的“次要影响”, 更能体现变量的动态特性.

### 2.2 mRMR-WDPCA 方法离线建模和在线监测

在上一节的动态特性描述方法的基础上, 我们给出如图 2 所示的包含离线和在线两部分的监控流程, 具体细节介绍如下.

#### 2.2.1 离线建模

**步骤 1.** 通过对  $X$  中每个测量变量引入其前  $d$  个采样值, 以获取如式 (2) 中的增广矩阵.

**步骤 2.** 对增广矩阵的每一个测量变量实施第 2.1 节的动态特性描述准则, 以获取  $X_a$  中每一列的权值向量, 并按式 (9) 求取  $X_i$ .

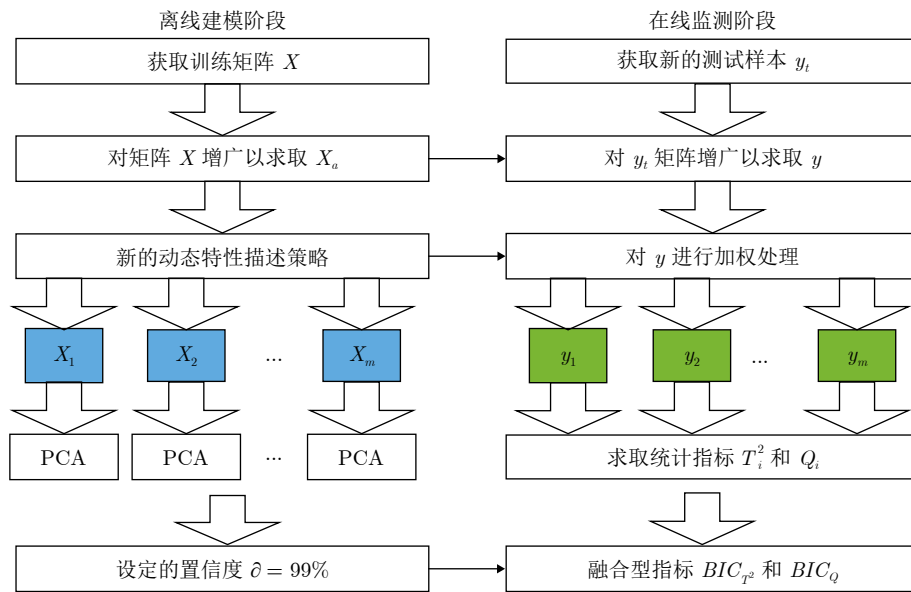


图 2 基于 mRMR-WDPCA 故障监测的流程图

Fig.2 Flowchart of mRMR-WDPCA based fault detection

步骤 3. 对  $X_i$  建立基于 PCA 的监测模型, 并重复此步骤以获得  $m$  个 PCA 监测模型.

### 2.2.2 在线监测

步骤 1. 对新的样本数据  $y_t \in \mathbf{R}^{m \times 1}$ , 先对其引入前  $d$  个时刻采样值以获取增广矩阵  $y = [y_t^T, y_{t-1}^T, \dots, y_{t-d}^T]^T$ .

步骤 2. 接着利用离线阶段得到的  $m$  个权值向量  $\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_m$ , 分别对  $y$  进行加权处理得到矩阵  $y_1, y_2, \dots, y_m$ , 并分别计算其在第  $i$  个 PCA 模型下的统计指标  $T_i^2$  和  $Q_i$ :

$$T_i^2 = y_i^T P_i \Lambda_i P_i^T y_i \quad (11)$$

$$Q_i = y_i^T (I - P_i P_i^T) y_i \quad (12)$$

对应的控制限的计算可参考文献 [21].

步骤 3. 采用了贝叶斯推理<sup>[17, 22]</sup> 将步骤 2 得到的  $m$  个子块的监测结果都转化为概率形式, 并将所有的概率值融合为一个综合性的概率型指标. 因此, 该指标非常适合于分布式过程监控. 以统计量  $T_i^2$  为例, 样本  $y_i$  的故障概率为

$$P_{T_i^2}(F|y_i) = \frac{P_{T_i^2}(y_i|F)P_{T_i^2}(F)}{P_{T_i^2}(y_i)} \quad (13)$$

其中, 概率  $P_{T_i^2}(y_i)$  定义为

$$P_{T_i^2}(y_i) = P_{T_i^2}(y_i|N)P_{T_i^2}(N) + P_{T_i^2}(y_i|F)P_{T_i^2}(F) \quad (14)$$

$N$  和  $F$  分别代表正常和故障,  $P_{T_i^2}(N), P_{T_i^2}(F)$

分别为  $\alpha$  和  $1 - \alpha$ . 条件概率  $P_{T_i^2}(y_i|N), P_{T_i^2}(y_i|F)$  可分别按下式计算:

$$P_{T_i^2}(y_i|N) = \exp\left(\frac{-T_i^2}{C_i}\right) \quad (15)$$

$$P_{T_i^2}(y_i|F) = \exp\left(\frac{-C_i}{T_i^2}\right) \quad (16)$$

其中,  $C_i$  为  $T_i^2$  统计量对应的控制限, 最后, 可按加权形式融合得到最终的概率监测指标  $BIC_{T^2}$

$$BIC_{T^2} = \sum_{i=1}^m \left( \frac{P_{T_i^2}(y_i|F)P_{T_i^2}(F|y_i)}{\sum_{i=1}^m P_{T_i^2}(y_i|F)} \right) \quad (17)$$

类似地, 可得到  $BIC_Q$

$$BIC_Q = \sum_{i=1}^m \left( \frac{P_{Q_i}(y_i|F)P_{Q_i}(F|y_i)}{\sum_{i=1}^m P_{Q_i}(y_i|F)} \right) \quad (18)$$

当  $BIC_{T^2} > 1 - \alpha$  或  $BIC_Q > 1 - \alpha$  时, 则认为样本  $y_t$  发生了故障, 否则认为过程处于正常工作状态.

## 3 仿真算例

TE (Tennessee-Eastman) 工业过程由文献 [23] 首次提出并已经成为验证过程控制和故障诊断方法的标准实验平台, 其流程如图 3 所示, 主要包含 5 个主要单元: 反应器、冷凝器、分离塔、汽提

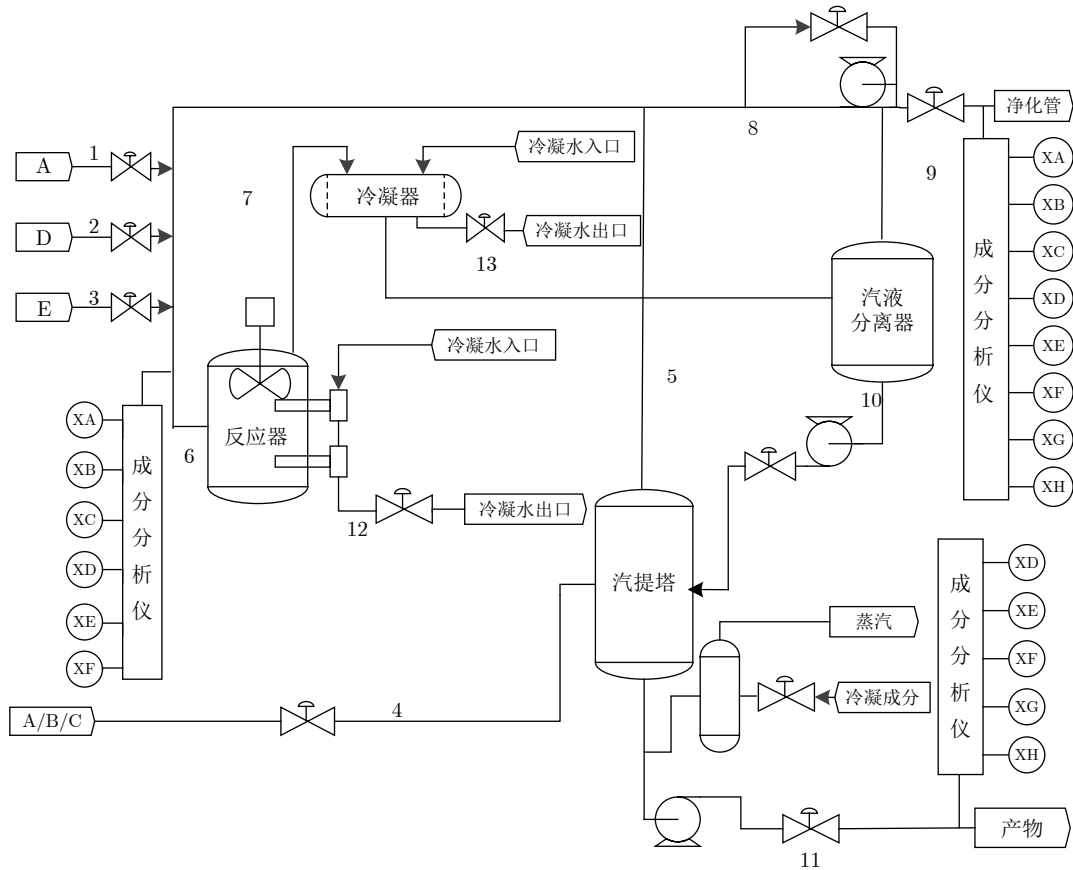


图 3 TE 过程的结构图

Fig.3 Structure diagram of the TE process

塔和压缩机. 主要涉及 8 种反应成分: A、B、C、D、E、F、G 和 H. 整个过程中共有 12 个操作变量和 41 个测量变量, 还有 21 种不同的预设故障, 每个故障都有相应的训练样本和测试样本<sup>[24]</sup>, 相应地, 故障都从第 160 个采样点开始引入, 且本文涉及的所有仿真数据都可从网站 <http://web.mit.edu/braatzgroup/links.html> 上下载. 本文依据文献 [25] 选择了过程中的 11 个操作变量和 22 个测量变量共 33 个变量作为监测对象, 且假设所有数据都服从高斯分布.

离线建模阶段, 将正常工况下的 960 数据作为训练集建立 DPCA、DLV 以及 MI-DPCA 的监控模型, 以作为本文方法的对比, 为了更好地进行比较, 所有方法的置信水平  $\alpha$  都设置为 99%. DLV 监测模型采用了 3 个统计量, 分别为  $T_d^2$ 、 $T_s^2$  和  $Q_r$ , 分别用于监测动态潜隐成分、静态潜隐成分和残差, 关于该模型的具体介绍可参考文献 [12], 此处不再赘述. 而对于 DPCA 和 MI-DPCA 模型中的延迟测量值  $d$  可根据文献 [7] 中提供的方法确定为 2, 保留的主元个数可通过累计方差贡献率准则 ( $CPV \geq$

85%) 确定为 34. 且本文提出的 mRMR-WDPCA 方法中的延迟测量值  $d = 2$ , 而每个子 PCA 模型中的主元个数同样由  $CPV \geq 85\%$  准则确定为 1.

首先利用 TE 过程中 500 组正常数据作为测试样本集测试 DPCA、DLV、MI-DPCA 以及本文的 mRMR-WDPCA 模型对正常工况的误报率 (第 I 型错误), 具体的实验结果如表 1 所示, 由表 1 可知, 虽然所提方法两个统计指标都没有取得最低的误报率, 但是从工程实践角度来说, 监控过程中连续 6 个采样点超过控制线时才被认为有故障发生, 而由于 mRMR-WDPCA 两个指标分别对应的误报率

表 1 TE 过程的误报率 (%)  
Table 1 False alarm rates of TE process (%)

模型	$T_s^2$	$BIC_{T_2}/T^2/T_d^2$	$BIC_Q/Q/Q_r$
DPCA	—	0.63	3.24
DLV	1.00	3.02	3.24
MI-DPCA	—	0.21	1.98
mRMR-WDPCA	—	1.63	2.13

1.63% 和 2.13% 都是通过单一采样点计算求得的, 所以所得结果仍然处于可接受的范围内, 所提方法在监控正常工况时也是有效的.

通常来讲, 较低的故障误报率会对应着较高的故障漏报率, 为了进一步说明所提方法的优越性, 还需对比 4 种方法的故障漏报率 (第 II 型错误). 即分别利用 4 种方法来检测 TE 过程的 21 种故障, 并记录每种方法不同统计量中较好的检测结果, 详见表 2. 其中, 由于故障 3, 9 和 15 发生时不会引起样本数据均值或方差的显著变化<sup>[26]</sup>, 因此很难被本文中以数据均值和方差为诊断特征的 mRMR-WDPCA 方法有效地检测出来, 所以, 上述 3 种故障在本研究中不予考虑. 但已有一些基于流形学习<sup>[27]</sup> 或子空间分解<sup>[28]</sup> 的数据驱动方法仍可以成功检测、诊断以上 3 种故障. 剩余的 18 种故障对应的最小的漏报率以及检测延迟数在表 2 中都标记为粗体, 以方便辨识.

从表 2 中可以看出, 所提方法在大多数故障类型上都能取得优越于其他三种方法的漏报率和检测延迟数. 特别地, 对于故障 10 和 16 的漏报率得到大幅的降低. 主要是因为 mRMR-WDPCA 模型考

虑变量间的线性和非线性关联同时减少变量间存在的冗余, 还通过不同权值的赋予, “凸显”了重要延迟变量作用的同时也“弱化”了次要延迟变量的次要影响, 因此所建立的监控模型能更加全面地描述当前测量变量的动态特性, 从而能进一步提高故障的监测效果. 即便是 mRMR-WDPCA 方法没有取得最好结果的故障, 但所得结果与最优的漏报率也相差无几. 此外, 4 种方法对剩余的 18 种故障的平均漏报率结果如图 4 所示, 其中,  $DLV_T$  表示 DLV 方法的两个统计量  $T_s^2$  和  $T_d^2$  的均值, 从图 4 中可知, 相比于其他三种方法, 本文所提模型的不同统计量都取得了最低的平均漏报率, 因此从另一个角度说明了方法的有效性和实用性.

为了更好地展示 mRMR-DDPCA 方法相比于其他对比方法的有效性, 特将 4 种方法对故障 10 和 16 这两种不同类型的故障的监控细节列于图 5 和图 6 中. 具体地, 从图 5(d) 中可以看出, mRMR-DDPCA 方法的  $BIC_Q$  统计量不但取得了最低的漏报率 18.88%, 而且在第 163 个采样点就检测出了故障, 故障延迟数为 2, 而其他三种方法对应的延迟数分别为 18, 7 和 24, 说明本文方法能在故障发生后

表 2 TE 过程故障漏报率 (%) 和检测延迟数 (个)  
Table 2 Missing alarm rates (%) and detection delay (delayed samples) of TE process

故障编号	故障类型	DPCA		DLV		MI-DPCA		mRMR-WDPCA	
		$T^2/Q$	检测延迟数	$T_s^2/T_d^2/Q_r$	检测延迟数	$BIC_{T2}/BIC_Q$	检测延迟数	$BIC_{T2}/BIC_Q$	检测延迟数
1	阶跃	0.13	<b>0</b>	<b>0.00</b>	<b>0</b>	0.13	<b>0</b>	0.25	<b>0</b>
2	阶跃	1.50	2	<b>1.00</b>	<b>0</b>	1.38	10	1.50	10
4	阶跃	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>
5	阶跃	55.00	<b>0</b>	0.13	<b>0</b>	73.13	<b>0</b>	<b>0.00</b>	<b>0</b>
6	阶跃	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>
7	阶跃	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>
8	随机	2.63	<b>1</b>	6.38	10	2.50	13	<b>1.75</b>	12
10	随机	48.88	18	37.50	7	25.50	24	<b>18.88</b>	<b>2</b>
11	随机	6.00	<b>3</b>	19.00	<b>3</b>	<b>4.63</b>	<b>3</b>	13.50	<b>3</b>
12	随机	0.88	<b>0</b>	9.00	<b>0</b>	0.63	<b>0</b>	<b>0.13</b>	<b>0</b>
13	慢偏移	<b>4.63</b>	35	4.88	<b>26</b>	<b>4.63</b>	39	5.38	41
14	粘滞	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>	<b>0.00</b>	<b>0</b>
16	未知	48.00	10	36.6	39	23.50	11	<b>14.75</b>	<b>7</b>
17	未知	2.25	16	5.13	16	<b>2.13</b>	<b>0</b>	3.38	<b>0</b>
18	未知	9.38	15	9.63	17	9.38	16	<b>9.00</b>	<b>1</b>
19	未知	<b>33.38</b>	<b>0</b>	37.00	10	37.63	1	65.002	2
20	未知	36.38	12	35.13	<b>2</b>	33.38	55	<b>32.50</b>	45
21	恒定故障	49.50	26	49.25	<b>7</b>	<b>42.63</b>	40	47.13	9

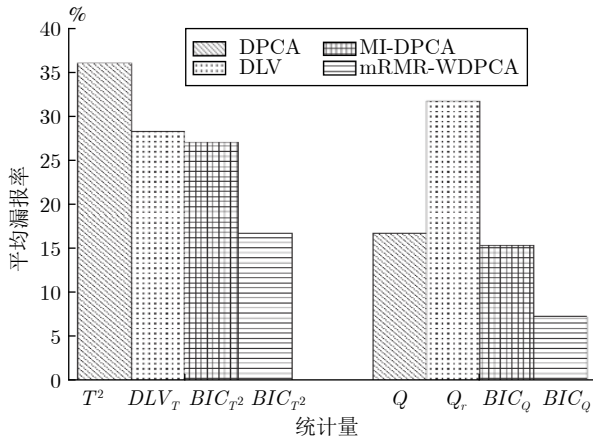


图 4 4 种方法的故障平均漏报率

Fig.4 Average missing alarm rates of the four methods

最快地将其检测出来. 此外, 当故障被检测到之后, mRMR-DDPCA 模型对应的统计量曲线很少回落到阈值线以下, 而 3 种对比算法的统计量曲线都会出现不同程度的回落现象, 从而导致较高的漏报率,

进而说明了所提方法故障检测的稳定性和持续性. 类似地, 图 6 中对于故障 16 的监测细节同样说明了所提方法的优越性.

### 4 结论

针对大型工业系统常常伴有复杂的动态特性, 且变量间的相互影响会体现在不同的采样时刻上等问题. 本文利用分布式 mRMR-WDPCA 方法充分刻画变量间相关性的同时, 也通过不同权值的赋予定量地描述了不同延迟变量对当前测量变量的影响程度, 凸显了变量间的相关性差异, 从而更好地解决了测量变量的动态特性问题, 取得较好的故障检测结果. 在 TE 过程上的仿真实验验证 mRMR-WDPCA 相对于其他方法的优越性. 然而, 本文仅局限于故障检测的分析, 而后续的故障诊断问题还未有所涉及. 此外, 所提的 mRMR-WDPCA 方法仍是一种线性模型且假设训练数据服从高斯分布, 如何将方法拓展以处理非线性、非高斯复杂过程仍

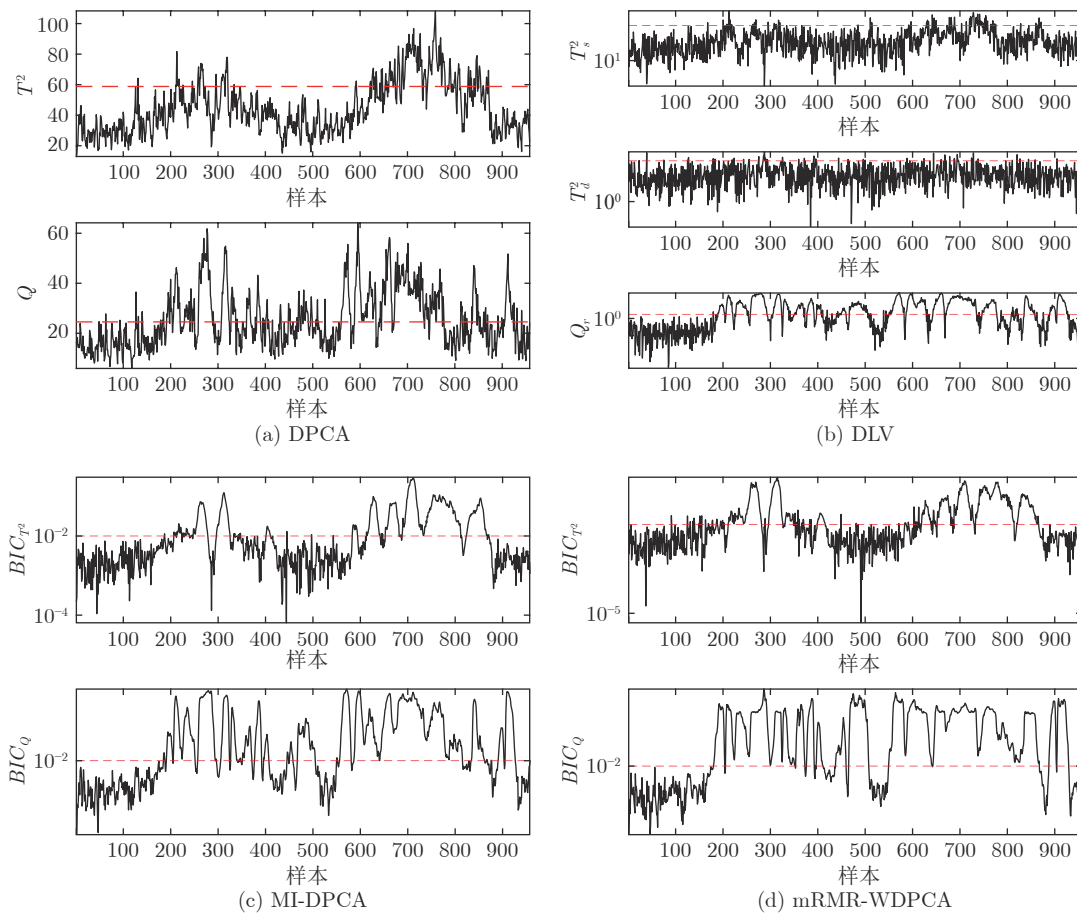


图 5 故障 10 的过程监控结果  
 Fig.5 The monitoring charts of Fault 10

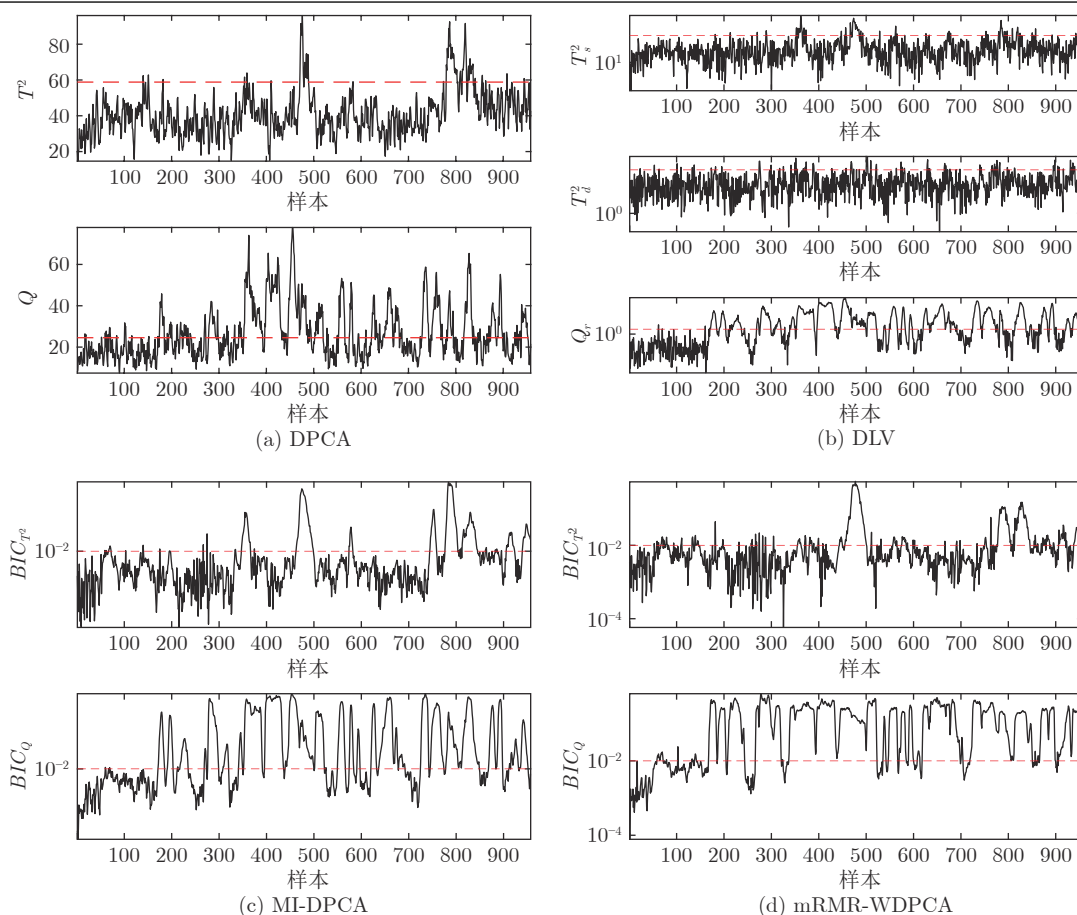


图 6 故障 16 的过程监控结果

Fig.6 The monitoring charts of Fault 16

是值得进一步研究之处。

## References

- Ge Z Q, Song Z H, Gao F R. Review of recent research on data-based process monitoring. *Industrial and Engineering Chemistry Research*, 2013, **52**(10): 3543–3562
- Peng Kai-Xiang, Ma Liang, Zhang Kai. Review of quality-related fault detection and diagnosis techniques for complex industrial processes. *Acta Automatica Sinica*, 2017, **43**(3): 349–365 (彭开香, 马亮, 张凯. 复杂工业过程质量相关的故障检测与诊断技术综述. *自动化学报*, 2017, **43**(3): 349–365)
- Yin S, Ding S X, Xie X, Luo H. A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics*, 2014, **61**(11): 6418–6428
- Jiang Q, Yan X. Nonlinear plant-wide process monitoring using MI-spectral clustering and Bayesian inference-based multiblock KPCA. *Journal of Process Control*, 2015, **32**: 38–50
- Jia Q L, Zhang Y. Quality-related fault detection approach based on dynamic kernel partial least squares. *Chemical Engineering Research and Design*, 2016, **106**: 242–252
- Han Min, Zhang Zhan-Kui. Fault detection method based on weighted kernel independent component analysis. *Control and Decision*, 2016, **31**(2): 242–248 (韩敏, 张占奎. 基于加权核独立成分分析的故障检测方法. *控制与决策*, 2016, **31**(2): 242–248)
- Dong Y N, Qin S J. A novel dynamic PCA algorithm for dynamic data modeling and process monitoring. *Journal of Process Control*, 2018, **67**: 1–11
- Zhou Dong-Hua, Hu Yan-Yan. Fault diagnosis techniques for dynamic systems. *Acta Automatica Sinica*, 2009, **35**(6): 748–758 (周东华, 胡艳艳. 动态系统的故障诊断技术. *自动化学报*, 2009, **35**(6): 748–758)
- Ku W, Storer R H, Georgakis C. Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 1995, **30**(1): 179–196
- Chen J, Liu K C. On-line batch process monitoring using dynamic PCA and dynamic PLS models. *Chemical Engineering Science*, 2002, **57**(1): 63–75
- Li W, Qin S J. Consistent dynamic PCA based on errors-in-variables subspace identification. *Journal of Process Control*, 2001, **11**(6): 661–678
- Li G, Qin S J, Zhou D H. A new method of dynamic latent-variable modeling for process monitoring. *IEEE Transactions on Industrial Electronics*, 2014, **61**(11): 6438–6445
- Zhang Y W, Zhou H, Qin S J. Decentralized fault diagnosis of large-scale processes using multiblock kernel principal component analysis. *Acta Automatica Sinica*, 2010, **36**(4): 593–597
- Ge Z Q, Song Z H. Distributed PCA model for plant-wide process monitoring. *Industrial and Engineering Chemistry Research*, 2013, **52**(5): 1947–1957
- Tong C D, Shi X H. Decentralized monitoring of dynamic processes based on dynamic feature selection and informative fault pattern dissimilarity. *IEEE Transactions on Industrial Electronics*, 2016, **63**(6): 3804–3814

- 16 Tong Chu-Dong, Lan Ting, Shi Xu-Hua. Fault detection by decentralized dynamic PCA algorithm on mutual information. *CIESC Journal*, 2016, **67**(10): 4317–4323  
(童楚东, 蓝艇, 史旭华. 基于互信息的分散式动态PCA故障检测方法. *化工学报*, 2016, **67**(10): 4317–4323)
- 17 Xu C, Zhao S Y, Liu F. Distributed plant-wide process monitoring based on PCA with minimal redundancy maximal relevance. *Chemometrics and Intelligent Laboratory Systems*, 2017, **169**: 53–63
- 18 Tong C D, L T, Shi X H. Fault detection and diagnosis of dynamic processes using weighted dynamic decentralized PCA approach. *Chemometrics and Intelligent Laboratory Systems*, 2017, **161**: 34–42
- 19 Han M, Zhong K, Qiu T, Han B. Interval type-2 fuzzy neural networks for chaotic time series prediction: A concise overview. *IEEE Transactions on Cybernetics*, 2019, **49**(7): 2720–2731
- 20 Peng H C, Long F H, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(8): 1226–1238
- 21 Qin S J. Statistical process monitoring: basics and beyond. *Journal of Chemometrics*, 2003, **17**(7–8): 480–502
- 22 Zhong K, Han M, Qiu T, Han B, Chen Y W. Distributed dynamic process monitoring based on minimal redundancy maximal relevance variable selection and Bayesian inference. *IEEE Transactions on Control Systems Technology*, 2020, **28**(5): 2037–2044
- 23 Downs J J, Vogel E F. A plant-wide industrial process control problem. *Computers and Chemical Engineering*, 1993, **17**(3): 245–255
- 24 Cao Yu-Ping, Huang Lin-Zhe, Tian Xue-Min. A process monitoring method using dynamic input-output canonical variate analysis. *Acta Automatica Sinica*, 2015, **41**(12): 2072–2080  
(曹玉苹, 黄琳哲, 田学民. 一种基于DIOCVCA的过程监控方法. *自动化学报*, 2015, **41**(12): 2072–2080)
- 25 Lee J M, Yoo C K, Lee I B. Statistical process monitoring with independent component analysis. *Journal of Process Control*, 2004, **14**(5): 467–485
- 26 Chiang L H, Russell E L, Braatz R D. *Fault Detection and Diagnosis in Industrial Systems*. London: Springer-Verlag, 2001.
- 27 Li N, Yan W W, Yang Y P. Spatial-statistical local approach for improved manifold-based process monitoring. *Industrial and Engineering Chemistry Research*, 2015, **54**(34): 8509–8519
- 28 Zhao C H, Wang W, Qin Y, Gao F R. Comprehensive subspace decomposition with analysis of between-mode relative changes for multimode process monitoring. *Industrial and Engineering Chemistry Research*, 2015, **54**(12): 3154–3166



**钟 凯** 大连理工大学电子信息与电气工程学部博士研究生. 主要研究方向为工业过程监控, 故障诊断.

E-mail: kaizhong0402@ahu.edu.cn

**(ZHONG Kai** Ph.D. candidate at the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology. His research interest covers industrial process monitoring and fault diagnosis.)



**韩 敏** 大连理工大学电子信息与电气工程学部教授. 主要研究方向为模式识别, 复杂系统建模与分析及时间序列预测. 本文通信作者.

E-mail: minhan@dlut.edu.cn

**(HAN Min** Professor at the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology. Her research interest covers pattern recognition, modeling and analysis of complex system, and time series prediction. Corresponding author of this paper.)



**韩 冰** 航运技术与安全国家重点实验室研究员. 主要研究方向为深海动力定位控制, 船舶动力装置的故障诊断和预测.

E-mail: hanbing@sssri.com

**(HAN Bing** Professor at the State Key Laboratory of Navigation and Safety Technology. His research interest covers deep sea dynamic positioning control, and fault diagnosis and prognostic of ship power plant.)