

# 基于质量关联虚拟变量的质量相关变量划分及故障检测

刘美枝<sup>1,2</sup> 孔祥玉<sup>2</sup> 胡昌华<sup>2</sup>

**摘要** 质量相关故障检测作为数据驱动的多元统计过程监测的重要研究内容,是保障复杂装备或工业过程安全高效运行的关键技术,而确定或划分质量相关变量是该方法的核心环节.现有质量相关故障检测方法通常高度依赖于质量变量,一旦质量变量不可测,其有效性便受到严重挑战.为解决这一挑战,提出基于质量关联虚拟变量(Quality-related virtual variable, QRV)的质量相关变量划分方法,基于此建立一种独立成分分析(Independent component analysis, ICA)质量相关故障检测模型,并开展故障检测应用研究.首先,构造一个 QRV,以间接反映系统的质量特性;其次,基于该 QRV,利用假设检验将过程变量划分为质量相关和质量无关变量组;随后,将该划分结果应用于基于 ICA 的质量相关故障检测,利用指数加权移动平均(Exponentially weighted moving average, EWMA)修正统计量,并构造综合检测指标;最后,通过数值仿真和田纳西-伊斯曼过程(Tennessee-Eastman process, TEP)实验验证了所提方法的可行性和有效性.

**关键词** 工业过程,质量相关变量,变量划分,质量关联虚拟变量,故障检测指标

**引用格式** 刘美枝,孔祥玉,胡昌华.基于质量关联虚拟变量的质量相关变量划分及故障检测.自动化学报,2025,51(5):1118-1130

**DOI** 10.16383/j.aas.c240627 **CSTR** 32138.14.j.aas.c240627

## Quality-related Variable Division and Fault Detection Based on Quality-related Virtual Variable

LIU Mei-Zhi<sup>1,2</sup> KONG Xiang-Yu<sup>2</sup> HU Chang-Hua<sup>2</sup>

**Abstract** Quality-related fault detection, as an important research content in data-driven multivariate statistical process monitoring, is a key technology for ensuring the safe and efficient operation of complex equipments or industrial processes. Determining or dividing quality-related variables is a core aspect of this method. Existing quality-related fault detection methods typically rely heavily on quality variables, and their effectiveness is severely challenged when quality variables are unmeasurable. To address this challenge, a quality-related variable division method based on quality-related virtual variable (QRV) is proposed. An independent component analysis (ICA)-based model for quality-related fault detection is established using this division method, and an application study on fault detection is conducted. First, a QRV is constructed to indirectly reflect the quality characteristics of systems; Second, based on this QRV, process variables are divided into quality-related and quality-unrelated variable groups by hypothesis testing; Subsequently, this division results are applied to the ICA-based quality-related fault detection, utilizing exponentially weighted moving average (EWMA) to modify statistics and construct comprehensive detection indices; Finally, the feasibility and effectiveness of the proposed method are verified through numerical simulations and Tennessee-Eastman process (TEP) experiments.

**Key words** Industrial processes, quality-related variable, variable division, quality-related virtual variable (QRV), fault detection indices

**Citation** Liu Mei-Zhi, Kong Xiang-Yu, Hu Chang-Hua. Quality-related variable division and fault detection based on quality-related virtual variable. *Acta Automatica Sinica*, 2025, 51(5): 1118-1130

收稿日期 2024-09-11 录用日期 2025-01-06

Manuscript received September 11, 2024; accepted January 6, 2025

国家自然科学基金(62273354, 62227814),山西省高等学校科技创新项目(2022L434)资助

Supported by National Natural Science Foundation of China (62273354, 62227814) and Scientific Innovation Foundation of the Higher Education Institutions of Shanxi Province (2022L434)

本文责任编辑 陈谋

Recommended by Associate Editor CHEN Mou

1. 山西大同大学物理与电子科学学院 大同 037009 2. 火箭军工程大学导弹工程学院 西安 710025

1. College of Physics and Electronic Science, Shanxi Datong University, Datong 037009 2. College of Missile Engineering, Rocket Force University of Engineering, Xi'an 710025

故障检测与诊断是保障复杂装备、工业过程安全稳定运行及产品质量的重要技术支撑<sup>[1]</sup>.相较于传统依赖机理模型或专家知识的方法,数据驱动的多元统计故障检测方法因其独特的优势而日益受到工程界和学术界的广泛关注,成为故障检测领域的研究热点<sup>[2]</sup>.该方法仅需利用系统运行时的离线和在线观测数据,借助较成熟的多元统计分析方法,特别适用于过程变量繁多且高度相关的系统.

质量相关故障检测是基于多元统计过程故障诊断的重要研究内容,而测量变量之间的相关性分析

是该方法的关键环节. 一般而言, 测量变量可分为两类: 过程变量和质量变量. 过程变量是指在复杂装备运行或工业生产过程中可以直接观测的物理量、化学量或操作参数, 如电压、电流、流量等, 这些变量反映装备运行或生产过程的实时状态. 质量变量是指与装备性能或生产质量指标相关的变量, 如火箭飞行姿态角, 轧钢或造纸的厚度、硬度等, 这些质量变量是评估装备性能或产品质量的依据.

需要指出的是, 由于复杂装备或工业过程结构复杂, 各组成部分之间高度耦合. 这些系统涉及的过程变量繁多, 且变量之间存在相关性. 但并非所有过程变量都与系统关键性能或产品质量密切相关. 在众多过程变量中, 通常只有部分变量的变化真正影响产品质量, 我们将这些变量称为质量相关变量, 而其余变量则被称为质量无关变量. 在故障检测中, 如果对所有过程变量进行建模, 质量无关变量的波动可能会导致故障检测指标超出控制限, 从而触发报警. 但这类故障并不直接影响产品质量, 我们称这类故障为质量无关故障. 由质量无关故障触发的报警不仅浪费资源和时间, 还可能误导操作人员对实际工况的判断, 大大影响生产效益. 为实现对生产过程的精准控制与优化, 确保产品质量和生产效益, 故障检测方法必须能够有效区分质量相关故障和质量无关故障.

一般情况下, 质量相关故障检测方法可以归纳为两大类: 一类为空间分解法, 另一类为变量划分法. 空间分解法, 也称为后分解法, 其主要思路是首先对所有过程变量建立模型, 形成投影空间, 随后将该投影空间分解为质量相关子空间和质量无关子空间. 这一类的典型方法包括全潜结构投影 (Total projection to latent structures, T-PLS)<sup>[3]</sup> 和改进偏最小二乘 (Modified partial least squares, MPLS)<sup>[4]</sup>, 这两种方法都建立在偏最小二乘 (Partial least squares, PLS) 模型的基础上, 对投影空间进行深入的分解和细化; 随后, 这两种方法的大量变体被相继提出<sup>[5-8]</sup>. 空间分解法不仅适用于 PLS 模型, 也被广泛应用于其他统计分析模型中, 如主成分分析 (Principal component analysis, PCA)<sup>[9]</sup>、典型相关分析 (Canonical correlation analysis, CCA)<sup>[10-11]</sup>、独立成分分析 (Independent component analysis, ICA)<sup>[12]</sup> 和慢特征分析 (Slow feature analysis, SFA)<sup>[13]</sup> 等. 例如, Wang 等<sup>[14]</sup> 利用核主成分分析提取核主成分, 随后基于该核主成分和质量变量建立 CCA 模型. Liu 等<sup>[15]</sup> 采用核独立成分对过程变量提取独立成分, 进一步建立独立成分与质量变量之间的回归模型, 提出修正的核独立成分回归 (Modi-

fied kernel independent component regression, MKICR) 模型. Zhang 等<sup>[16]</sup> 提出一种基于质量相关修正正则化的 SFA 分布式故障检测方法, 以处理大规模时序过程中的动态、连接关系及异常值问题.

变量划分法根据过程变量与质量变量的相关性强弱, 将过程变量分为质量相关组和质量无关组. 随后, 对这两组变量分别建模实现质量相关故障检测. 典型的划分方法包括互信息 (Mutual information, MI)<sup>[17-18]</sup>, 最小冗余最大相关 (Minimal redundancy maximal relevance, mRMR)<sup>[19]</sup>, 传递熵 (Transfer entropy, TE)<sup>[20-21]</sup>, 投影变量重要性 (Variable importance in the projection, VIP)<sup>[22-23]</sup> 等. 例如, Guo 等<sup>[24]</sup> 结合先验知识和测量变量建立时空拓扑图, 并利用 MI 划分变量构建质量相关子空间, 进行质量相关故障检测. Zhu 等<sup>[25]</sup> 利用 MI 选择质量相关变量, 随后根据操作单元将整个过程分解为多个块, 分别在每个块中采用改进的卷积神经网络建立特征提取模型. Yu 等<sup>[26]</sup> 利用 TE 计算变量间的因果关系, 对质量相关变量提取卷积特征, 建立质量相关故障检测模型.

上述空间分解法和变量划分法的应用均存在一个前提条件: 质量变量必须可测, 因为它们依赖质量变量来分解质量相关子空间或划分质量相关变量. 然而, 在实际工业过程中, 由于生产环境恶劣或测量技术的局限性, 某些关键的质量变量, 例如脱丁烷塔塔底丁烷浓度, 可能难以甚至无法通过现有的测量仪器准确测量. 这一情况对传统的质量相关故障检测方法构成严峻的挑战. 针对这一挑战, 本文提出一种基于质量关联虚拟变量 (Quality-related virtual variable, QRV) 的质量相关变量划分方法, 建立一种基于 QRV 的 ICA 质量相关故障检测模型 (QRV-based ICA, QRVICA). 值得注意的是, QRV 的构造过程仅使用过程变量, 不依赖质量变量. 然而, 该 QRV 仍保持与质量变量较强的相关性, 并能够有效追踪质量变量的变化趋势. 因此, QRV 可作为一种虚拟量间接反映系统的质量特性.

与现有质量相关故障检测方法相比, 本文的主要贡献可归纳为: 1) 针对工业过程中质量变量不可测的情况, 提出一种质量关联虚拟变量构造方法. 该虚拟变量可间接反映系统的质量特性, 为变量间的相关性分析提供一种新方案. 2) 基于所构造的 QRV, 提出一种基于假设检验和滑动窗口的质量相关变量划分方法, 为后续故障检测奠定基础, 有效克服传统质量相关故障检测对质量变量的直接依赖. 3) 设计一套完整的故障检测策略, 通过指数加权移动平均 (Exponentially weighted moving aver-

age, EWMA) 将历史样本信息融入当前样本的检测指标构造中, 实现兼具最低误报率和最高检测率的质量相关故障检测。

## 1 基于 ICA 的故障检测

本节简要介绍基于 ICA 的故障检测. 给定标准化训练数据矩阵  $X \in \mathbf{R}^{n \times m}$ , 其中  $n$  和  $m$  分别表示样本数和变量数. ICA 的核心思想是  $m$  维的测量数据可以表示为  $d$  维 ( $d \leq m$ ) 独立成分的线性组合:

$$X = SA + E \quad (1)$$

式中,  $S = [s_1, \dots, s_d]$  为独立成分矩阵;  $A$  和  $E$  分别为混合矩阵和残差矩阵. ICA 的关键是由已知数据矩阵  $X$  估计混合矩阵  $A$  和独立成分矩阵  $S$ , 该问题可转化为寻求混解矩阵  $W$ , 使得  $S = XW$  中各独立分量彼此独立. 为消除变量间的相关性, 对矩阵  $X$  进行白化. 首先对  $X$  的协方差矩阵  $S_F$  进行特征值分解:

$$S_F = \frac{1}{n-1} X^T X = U \Lambda U^T \quad (2)$$

其中,  $\Lambda$  和  $U$  分别为特征值矩阵和特征向量矩阵. 经白化后的数据矩阵可表示为:

$$Z = XR \approx SAR = SC \quad (3)$$

式中  $R = U\Lambda^{-1/2}$  为白化矩阵;  $C = AR$  为正交矩阵. 这样, 独立成分矩阵  $S$  的估计可转化为:

$$S = ZC^T = XRC^T \quad (4)$$

至此, 混解矩阵  $W$  可表示为:

$$W = RC^T \quad (5)$$

通常采用快速 ICA (FastICA) 求解上述混解矩阵, 以最大化负熵为准则函数, 负熵采用非二次函数  $G(\cdot)$  近似, 则目标函数可表示为:

$$J(c_i) = \max_{c_i \in C} [E\{G(Zc_i^T)\} - E\{G(v)\}]^2 \quad (6)$$

s.t.  $E\{(Zc_i^T)^2\} = 1$

式中,  $E\{\cdot\}$  代表期望;  $Zc_i^T = s_i$  是第  $i$  个独立成分;  $v$  表示零均值单位方差的高斯随机变量. 具体推导和求解过程可参考文献 [26].

为监测系统运行状态, 通常采用  $I^2$  统计量监测独立成分空间的变化, 用  $Q$  统计量监测残差空间的变化. 对于样本  $x$ , 其统计量可计算如下:

$$I^2(x) = ss^T \quad (7)$$

$$Q(x) = ee^T \quad (8)$$

式中,  $s = xW$  为样本  $x$  的独立成分得分;  $e = x -$

$sA$  为样本残差. 考虑到  $s$  和  $e$  可能无法满足高斯分布, 采用核密度估计 (Kernel density estimation, KDE) 计算其相应的控制限  $I_{\text{lim}}^2$  和  $Q_{\text{lim}}$ .

## 2 质量相关变量划分方法

在复杂装备或工业过程中, 由于系统内部物料、能量及信息间的传递和反馈控制的存在, 质量相关变量之间形成错综复杂的相互作用和依赖关系<sup>[2]</sup>. 因此, 这些质量相关变量在系统运行过程中往往表现出协同变化的趋势, 而质量无关变量的变化趋势可能与质量相关变量缺乏明显的关联性. 当质量变量难以直接测量时, 是否存在一个虚拟变量 (我们称之为“质量关联虚拟变量”, QRV), 该变量能够与所有的质量相关变量均展现出明显的相关性, 而那些质量无关变量则没有显著关联? 如果这样的 QRV 确实存在, 它便可作为表征系统质量特性的综合指标, 有效克服质量相关变量划分对质量变量的依赖. 这样, 通过评估各过程变量与 QRV 之间的相关性强度, 我们便可准确地划分质量相关变量. 按照上述思路, 本文所提基于 QRV 的质量相关变量划分方法的整体框架如图 1 所示.

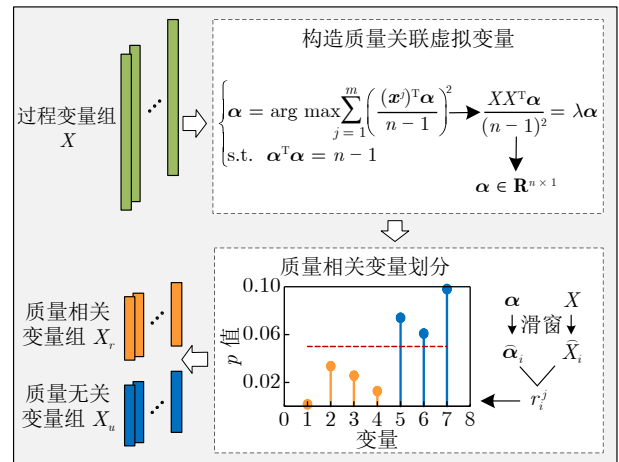


图 1 基于 QRV 的质量相关变量划分方法的整体框架  
Fig. 1 Overall framework of the quality-related variable division method based on QRV

### 2.1 质量关联虚拟变量的构造

给定复杂系统正常工况下的过程变量矩阵  $X \in \mathbf{R}^{n \times m}$  如下:

$$X = [x_1; \dots; x_n] = [x^1, \dots, x^m] \quad (9)$$

式中  $x_i$  表示  $i$  时刻的采样样本;  $x^j$  表示第  $j$  个过程变量的时序数据. 假设  $X$  已经标准化 (即具有零均值单位方差). 构造质量关联虚拟变量  $\alpha \in \mathbf{R}^{n \times 1}$ , 目

标函数为最大化其与所有过程变量的内积平方和:

$$\begin{aligned} \boldsymbol{\alpha} &= \arg \max \sum_{j=1}^m \left( \frac{(\boldsymbol{x}^j)^T \boldsymbol{\alpha}}{n-1} \right)^2 \\ \text{s.t. } \boldsymbol{\alpha}^T \boldsymbol{\alpha} &= n-1 \end{aligned} \quad (10)$$

利用拉格朗日乘子法, 上述目标函数可转化为:

$$L(\boldsymbol{\alpha}) = \sum_{j=1}^m \left( \frac{(\boldsymbol{x}^j)^T \boldsymbol{\alpha}}{n-1} \right)^2 + \lambda(n-1 - \boldsymbol{\alpha}^T \boldsymbol{\alpha}) \quad (11)$$

式中,  $\lambda$  为拉格朗日算子, 对  $L(\boldsymbol{\alpha})$  求导可得:

$$\begin{aligned} \frac{\partial L(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} &= \frac{2}{(n-1)^2} \sum_{j=1}^m (\boldsymbol{x}^j)^T \boldsymbol{\alpha} \boldsymbol{x}^j - 2\lambda \boldsymbol{\alpha} = \\ &= \frac{2 [(\boldsymbol{x}^1)^T \boldsymbol{\alpha} \boldsymbol{x}^1 + \cdots + (\boldsymbol{x}^m)^T \boldsymbol{\alpha} \boldsymbol{x}^m]}{(n-1)^2} - 2\lambda \boldsymbol{\alpha} \end{aligned} \quad (12)$$

令  $\frac{\partial L(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} = 0$ , 对式 (12) 进行化简, 式 (10) 优化问题可转化为如下特征值分解问题:

$$\frac{X X^T \boldsymbol{\alpha}}{(n-1)^2} = \lambda \boldsymbol{\alpha} \quad (13)$$

选择最大特征值对应的特征向量  $\boldsymbol{\alpha}$  作为 QRV.

## 2.2 质量相关变量的划分

基于所构造 QRV, 将那些与 QRV 相关度高的过程变量确定为质量相关变量, 而将与 QRV 相关度低的过程变量确定为质量无关变量. 为综合考虑变量对之间的局部相关性, 采用窗口宽度为  $M$  的滑窗将时序数据截取为  $n-M+1$  个子块, 如图 2 所示, 图中第  $i$  个子块记作:

$$\begin{aligned} \widehat{X}_i &= [\boldsymbol{x}_i; \boldsymbol{x}_{i+1}; \cdots; \boldsymbol{x}_{i+M-1}] = \\ &= [\widehat{\boldsymbol{x}}_i^1, \widehat{\boldsymbol{x}}_i^2, \cdots, \widehat{\boldsymbol{x}}_i^m] \end{aligned} \quad (14)$$

同理, QRV 的第  $i$  个子块记作:

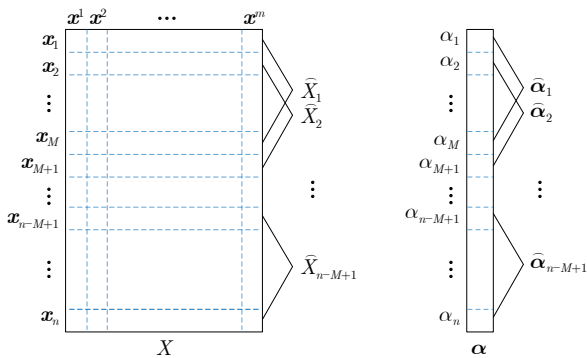


图 2 时间序列分块示意图

Fig.2 Schematic diagram of time series chunking

$$\widehat{\boldsymbol{\alpha}}_i = [\alpha_i, \alpha_{i+1}, \cdots, \alpha_{i+M-1}]^T \quad (15)$$

由  $\widehat{X}_i$  和  $\widehat{\boldsymbol{\alpha}}_i$  ( $i \in [1, n-M+1]$ ) 组成子块对, 在每个子块对中计算各过程变量与 QRV 的相关度, 最后利用假设检验确定过程变量是否为质量相关变量. 在第  $i$  个子块对  $\widehat{X}_i$  和  $\widehat{\boldsymbol{\alpha}}_i$  ( $i \in [1, n-M+1]$ ) 中, 第  $j$  个变量  $\widehat{\boldsymbol{x}}_i^j$  与  $\widehat{\boldsymbol{\alpha}}_i$  的相关系数为:

$$r_i^j = \frac{\sum_{k=1}^M (\widehat{x}_{i,k}^j - \bar{x}_i^j) (\widehat{\alpha}_{i,k} - \bar{\alpha}_i)}{\sqrt{\sum_{k=1}^M (\widehat{x}_{i,k}^j - \bar{x}_i^j)^2} \sqrt{\sum_{k=1}^M (\widehat{\alpha}_{i,k} - \bar{\alpha}_i)^2}} \quad (16)$$

式中,  $\widehat{x}_{i,k}^j$  表示向量  $\widehat{\boldsymbol{x}}_i^j$  的第  $k$  个元素;  $\bar{x}_i^j$  表示向量  $\widehat{\boldsymbol{x}}_i^j$  的均值;  $\widehat{\alpha}_{i,k}$  表示向量  $\widehat{\boldsymbol{\alpha}}_i$  的第  $k$  个元素;  $\bar{\alpha}_i$  表示向量  $\widehat{\boldsymbol{\alpha}}_i$  的均值. 构造  $t$  统计量:

$$t_i^j = r_i^j \sqrt{\frac{M-2}{1-(r_i^j)^2}} \quad (17)$$

取原假设和备择假设分别为:

$$\begin{cases} H_0: r_i^j = 0, \text{ 即两变量相关度为 } 0 \\ H_1: r_i^j \neq 0, \text{ 即两变量相关度不为 } 0 \end{cases} \quad (18)$$

采用双尾检验计算  $r_i^j$  对应的  $p$  值, 记作  $p_i^j$ . 分别计算  $n-M+1$  个子块的  $p$  值并组成检验矩阵:

$$P = \begin{bmatrix} p_1^1 & \cdots & p_1^m \\ p_2^1 & \cdots & p_2^m \\ \vdots & \ddots & \vdots \\ p_{n-M+1}^1 & \cdots & p_{n-M+1}^m \end{bmatrix} \in \mathbf{R}^{(n-M+1) \times m} \quad (19)$$

对矩阵  $P$  按列求均值  $\bar{P} = [\bar{p}^1, \cdots, \bar{p}^m]$ , 根据上述假设, 取显著性水平为 0.05. 如果  $\bar{p}^j < 0.05$ , 则原假设  $H_0$  不成立, 即过程变量  $\boldsymbol{x}^j$  为质量相关变量, 否则为质量无关变量.

至此, 过程变量  $X$  可被划分为质量相关变量组  $X_r$  和质量无关变量组  $X_u$ .

## 3 基于 QRV 的质量相关故障检测

本节在前述变量划分结果的基础上, 详述基于 ICA 的故障检测指标构建和故障检测策略.

按照式 (1) ~ (6), 分别对质量相关变量组  $X_r$  和质量无关变量组  $X_u$  建立 ICA 模型:

$$\begin{cases} X_r = S_r A_r + E_r \\ X_u = S_u A_u + E_u \end{cases} \quad (20)$$

$$\begin{cases} S_r = X_r W_r \\ S_u = X_u W_u \end{cases} \quad (21)$$

式中,  $A_r$  和  $A_u$  为混合矩阵;  $W_r$  和  $W_u$  为混解矩阵;

$S_r$  和  $S_u$  为独立成分;  $E_r$  和  $E_u$  则表示残差; 下标  $r$  和  $u$  分别表示质量相关空间和质量无关空间.

### 3.1 综合检测指标的构建

对于  $t$  时刻样本  $\mathbf{x}_t$ , 由质量相关变量和质量无关变量组成的向量分别记作  $\mathbf{x}_{t,r}$  和  $\mathbf{x}_{t,u}$ . 则样本  $\mathbf{x}_t$  的质量相关统计量按照式 (7) 和 (8) 可表示为:

$$I_r^2(\mathbf{x}_t) = \mathbf{s}_{t,r} \mathbf{s}_{t,r}^T \quad (22)$$

$$Q_r(\mathbf{x}_t) = \mathbf{e}_{t,r} \mathbf{e}_{t,r}^T \quad (23)$$

式中  $\mathbf{s}_{t,r} = \mathbf{x}_{t,r} W_r$  和  $\mathbf{e}_{t,r} = \mathbf{x}_{t,r} - \mathbf{s}_{t,r} A_r$  分别表示  $\mathbf{x}_t$  在质量相关空间的独立成分和残差. 同理可求得  $\mathbf{x}_t$  的质量无关统计量  $I_u^2(\mathbf{x}_t)$  和  $Q_u(\mathbf{x}_t)$ . 采用 KDE 分别求得  $I_r^2(\mathbf{x}_t)$ ,  $Q_r(\mathbf{x}_t)$ ,  $I_u^2(\mathbf{x}_t)$  和  $Q_u(\mathbf{x}_t)$  的控制限, 记作  $I_{r\_lim}^2$ ,  $Q_{r\_lim}$ ,  $I_{u\_lim}^2$  和  $Q_{u\_lim}$ .

复杂系统运行中, 由于环境变化、电磁干扰, 测量变量会发生波动. 为克服这种波动给检测结果带来的负面影响, 本文采用 EWMA 对统计量进行修正:

$$\begin{cases} \widehat{I}_r^2(\mathbf{x}_t) = \gamma I_r^2(\mathbf{x}_t) + (1-\gamma) \widehat{I}_r^2(\mathbf{x}_{t-1}) \\ \widehat{Q}_r(\mathbf{x}_t) = \gamma Q_r(\mathbf{x}_t) + (1-\gamma) \widehat{Q}_r(\mathbf{x}_{t-1}) \end{cases} \quad (24)$$

$$\begin{cases} \widehat{I}_u^2(\mathbf{x}_t) = \gamma I_u^2(\mathbf{x}_t) + (1-\gamma) \widehat{I}_u^2(\mathbf{x}_{t-1}) \\ \widehat{Q}_u(\mathbf{x}_t) = \gamma Q_u(\mathbf{x}_t) + (1-\gamma) \widehat{Q}_u(\mathbf{x}_{t-1}) \end{cases} \quad (25)$$

式中,  $\gamma \in [0, 1]$  为遗忘因子, 其值越小表示记忆深度越深. 初值  $\widehat{I}_r^2(\mathbf{x}_0)$ ,  $\widehat{Q}_r(\mathbf{x}_0)$ ,  $\widehat{I}_u^2(\mathbf{x}_0)$  和  $\widehat{Q}_u(\mathbf{x}_0)$  用全部训练样本统计量的均值代替.

为简化诊断逻辑, 用统计量的最大变化率来构建综合检测指标:

$$\varphi_r(\mathbf{x}_t) = \max \left[ \frac{\widehat{I}_r^2(\mathbf{x}_t)}{I_{r\_lim}^2}, \frac{\widehat{Q}_r(\mathbf{x}_t)}{Q_{r\_lim}} \right] \quad (26)$$

$$\varphi_u(\mathbf{x}_t) = \max \left[ \frac{\widehat{I}_u^2(\mathbf{x}_t)}{I_{u\_lim}^2}, \frac{\widehat{Q}_u(\mathbf{x}_t)}{Q_{u\_lim}} \right] \quad (27)$$

式中,  $\varphi_r(\mathbf{x}_t)$  和  $\varphi_u(\mathbf{x}_t)$  分别表示质量相关和质量无关空间的综合检测指标, 它们的控制限均为 1.

### 3.2 故障检测策略

为方便理解所提 QRVICA 质量相关故障检测方法, 详细的离线建模和在线检测步骤总结如下:

离线建模:

1) 将正常工况过程变量矩阵  $X$  标准化为零均值单位方差;

2) 按照式 (13) 进行特征值分解, 求解得质量

关联虚拟变量  $\alpha$ ;

3) 按照图 2 和式 (14) ~ (15) 对  $X$  和  $\alpha$  进行分块;

4) 按照式 (16) 分别计算每个子块对中各变量  $\widehat{\mathbf{x}}_i^j$  与  $\widehat{\alpha}_i$  的相关系数  $r_i^j$ ;

5) 按照式 (17) 计算  $t$  统计量  $t_i^j$ , 并利用双尾检验求得对应的  $p$  值  $p_i^j$ ;

6) 对式 (19) 所示矩阵  $P$  按列求均值向量  $\bar{P} = [\bar{p}^1, \bar{p}^2, \dots, \bar{p}^m]$ ;

7) 利用假设检验, 将  $X$  划分为质量相关变量组  $X_r$  和质量无关变量组  $X_u$ ;

8) 分别对  $X_r$  和  $X_u$  建立 ICA 模型, 计算控制限. 在线检测:

1) 对  $t$  时刻样本  $\mathbf{x}_t$  进行标准化;

2) 根据变量划分结果, 将  $\mathbf{x}_t$  划分为  $\mathbf{x}_{t,r}$  和  $\mathbf{x}_{t,u}$ ;

3) 将  $\mathbf{x}_{t,r}$  和  $\mathbf{x}_{t,u}$  分别投影到 ICA 模型中, 并求得相应的统计量  $\widehat{I}_r^2(\mathbf{x}_t)$ ,  $\widehat{Q}_r(\mathbf{x}_t)$ ,  $\widehat{I}_u^2(\mathbf{x}_t)$  和  $\widehat{Q}_u(\mathbf{x}_t)$ ;

4) 按照式 (26) ~ (27) 求得综合检测指标  $\varphi_r(\mathbf{x}_t)$  和  $\varphi_u(\mathbf{x}_t)$ ;

5) 如果  $\varphi_r(\mathbf{x}_t) > 1$ ,  $\mathbf{x}_t$  为质量相关故障样本; 如果  $\varphi_r(\mathbf{x}_t) \leq 1$  且  $\varphi_u(\mathbf{x}_t) > 1$ ,  $\mathbf{x}_t$  为质量无关故障样本; 如果  $\varphi_r(\mathbf{x}_t) \leq 1$  且  $\varphi_u(\mathbf{x}_t) \leq 1$ ,  $\mathbf{x}_t$  为正常样本.

基于 QRVICA 的质量相关故障检测方法流程如图 3 所示.

## 4 实验仿真

本节通过数值仿真和田纳西-伊斯曼过程 (Tennessee-Eastman process, TEP) 工业基准案例验证所提方法的有效性. 实验主要包括质量相关变量划分、相关性分析和故障检测. 在质量相关变量划分实验中, 选择互信息法<sup>[17]</sup> 作为对比方法. 在故障检测实验中, 为全面评估所提方法的故障检测性能, 我们选择 4 种质量相关故障检测方法做对比实验, 其中 MI-KPCA<sup>[17]</sup> 为基于变量划分的方法, MKICR<sup>[15]</sup> 和 OMDPLS<sup>[7]</sup> 为基于空间分解的方法, 而 VIP-DCPLS<sup>[22]</sup> 则结合变量划分和空间分解法. 此外, 为验证 EWMA 在所提 QRVICA 方法中的作用, 我们进行消融实验, 即去除 QRVICA 方法中的 EWMA 部分 (设置遗忘因子  $\gamma = 1$ ), 记为 QRVICA-without-EWMA. 文献 [22] 中, VIP-DCPLS 通过支持向量数据描述 (Support vector data description, SVDD) 整合质量相关和质量无关空间的 10 个统计量, 形成综合检测指标  $d$ , 但该指标  $d$  无法区分故障是否与质量相关. 为方便性能对比, 本实验中对 VIP-DCPLS 进行简单修改: 分别用 SVDD 处理质量相关空间的 5 个统计量得到质量相关指标  $d_r$ , 对质量无关空间的 5 个统计量构建质

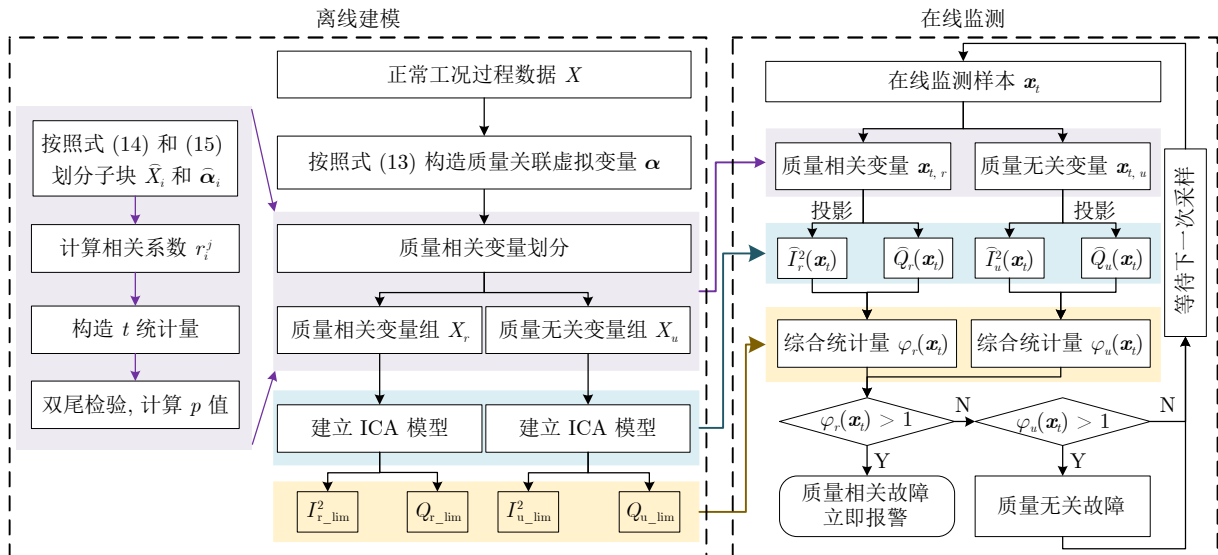


图3 基于QRVICA的质量相关故障检测流程图

Fig.3 Flowchart of quality-related fault detection based on QRVICA

量无关指标  $d_u$ . 所有方法中控制限的置信限设置为 0.95<sup>[15]</sup>, 根据交叉验证设置遗忘因子  $\gamma = 0.2$ . 采用误报率 (False alarm rate, FAR) 和故障检测率 (Fault detection rate, FDR) 定量评估各方法的检测性能. 具体计算公式如下:

FAR =

$$\begin{cases} \frac{\text{Num}\{\varphi_r(\mathbf{x}) > 1 | \mathbf{x} \in \text{Normal}\}}{\text{Num}\{\mathbf{x} \in \text{Normal}\}}, & \text{质量相关故障} \\ \frac{\text{Num}\{\varphi_u(\mathbf{x}) > 1\}}{\text{Num}\{\mathbf{x}\}}, & \text{质量无关故障} \end{cases} \quad (28)$$

FDR =

$$\begin{cases} \frac{\text{Num}\{\varphi_r(\mathbf{x}) > 1 | \mathbf{x} \in \text{Fault}\}}{\text{Num}\{\mathbf{x} \in \text{Fault}\}}, & \text{质量相关故障} \\ \frac{\text{Num}\{\varphi_u(\mathbf{x}) > 1 | \mathbf{x} \in \text{Fault}\}}{\text{Num}\{\mathbf{x} \in \text{Fault}\}}, & \text{质量无关故障} \end{cases} \quad (29)$$

式中,  $\text{Num}\{\varphi_r(\mathbf{x}) > 1 | \mathbf{x} \in \text{Normal}\}$  是  $\varphi_r(\mathbf{x}) > 1$  的正常样本个数;  $\text{Num}\{\varphi_r(\mathbf{x}) > 1 | \mathbf{x} \in \text{Fault}\}$  是  $\varphi_r(\mathbf{x}) > 1$  的故障样本的个数;  $\text{Num}\{\varphi_u(\mathbf{x}) > 1 | \mathbf{x} \in \text{Fault}\}$  表示  $\varphi_u(\mathbf{x}) > 1$  的故障样本的个数;  $\text{Num}\{\mathbf{x}\}$  表示所有样本 (包括正常样本和故障样本) 的个数;  $\text{Num}\{\mathbf{x} \in \text{Normal}\}$  和  $\text{Num}\{\mathbf{x} \in \text{Fault}\}$  则分别表示正常和故障样本的个数.

一般情况下 FAR 越低越好, FDR 越高越好. 在质量相关故障检测中, 注意以下两点:

1) 对于质量相关故障, 聚焦于质量相关检测指标. 理想的检测方法的质量相关检测指标需大于对

应的控制限, 从而确保对质量相关故障的敏感捕捉与及时响应, 即高 FDR.

2) 对于质量无关故障, 需同时评估质量相关与质量无关两个检测指标. 理想的检测方法应实现以下平衡: 一方面, 质量相关检测指标大于控制限的情形属于误报警, 质量相关检测指标需保持在控制限以下, 以最小化误报, 即低 FAR; 另一方面, 质量无关检测指标需高于控制限, 确保对非质量因素导致的潜在问题亦能进行有效预警, 即高 FDR.

#### 4.1 数值仿真

按照式 (30) 建立数值仿真模型<sup>⑧</sup>:

$$\begin{cases} \mathbf{x}_t = \mathbf{z}_t A + \mathbf{e}_t \\ \mathbf{y}_t = \mathbf{x}_t C + \mathbf{v}_t \end{cases} \quad (30)$$

式中,  $A = \begin{bmatrix} 1 & 3 & 4 & 4 & 0 & 0 & 0 \\ 3 & 1 & 1 & 4 & 1 & 1 & 1 \\ 1 & 1 & 3 & 0 & 0 & 2 & 3 \end{bmatrix}$ ;  $C = [2 \ 3 \ 1 \ 1 \ 0 \ 0 \ 0]^T$ ;  $\mathbf{x}_t$  为过程变量;  $\mathbf{y}_t$  为质量变量;  $\mathbf{z}_t \sim U(0, 1)$ ;  $\mathbf{e}_t \sim N(0, 0.05^2)$  和  $\mathbf{v}_t \sim N(0, 0.01^2)$  表示随机噪声.  $\mathbf{x}^1 \sim \mathbf{x}^4$  为质量相关变量, 而  $\mathbf{x}^5 \sim \mathbf{x}^7$  为质量无关变量.

##### 4.1.1 质量相关变量划分实验

按照式 (30) 生成 500 个正常样本, 采用互信息法<sup>[17]</sup> 作为对比方法验证所提方法的有效性, 实验结果如图 4. 图中用橙色标记质量相关变量, 用蓝色标记质量无关变量. 在互信息法中, 计算各过程变量与式 (30) 中质量变量  $\mathbf{y}_t$  之间的互信息值, 图 4(a) 中红色虚线表示控制限, 当某一变量与  $\mathbf{y}_t$  的互信息值

高于控制限, 该变量被划分为质量相关变量; 在所提方法中设置窗口宽度  $M = 10$ , 按照第 2.2 节计算所构造 QRV 与各过程变量之间的相关度, 图 4(b) 中红色虚线为显著性水平. 按照式 (18) 假设, 当某一变量的  $p$  值低于显著性水平, 则该变量为质量相关变量.

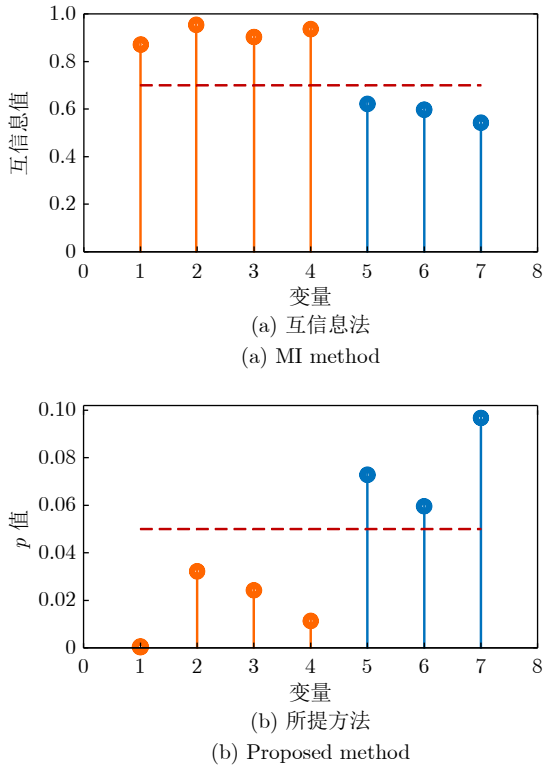


图 4 数值仿真过程变量划分结果  
Fig.4 Division results of process variable in numerical simulation

由图 4 可知, 所提基于 QRV 的质量相关变量划分方法与互信息法的实验结果一致, 均表明  $x^1 \sim x^4$  为质量相关变量, 而  $x^5 \sim x^7$  为质量无关变量, 该结果与实际情况一致. 该一致性说明, 即便在不使用质量变量的情况下, 利用本文所提基于 QRV 的划分方法, 仍能准确地将过程变量划分为质量相关变量和质量无关变量.

4.1.2 相关性分析实验

为进一步阐释所构造 QRV 可反映质量特性, 我们采用 4 阶 Butterworth 低通滤波器分别提取式 (30) 中质量变量  $y_t$  和质量关联虚拟变量  $\alpha$  的包络线如图 5 所示, 以此来直观查看信号随时间的变化趋势. 由图 5 可知, 两个信号的包络线在时间轴上呈现出近乎一致的变化轨迹, 这意味着 QRV 可有效跟踪  $y_t$  的变化.

此外, 为定量分析质量变量  $y_t$  及质量关联虚拟

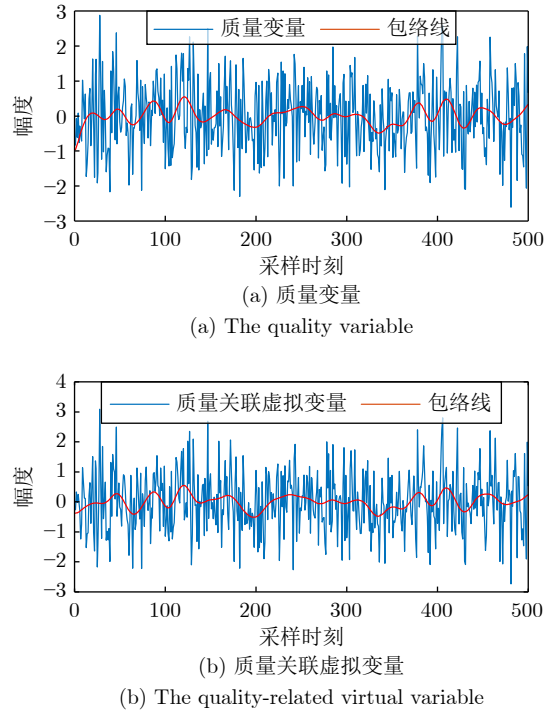


图 5 变量的变化趋势及其包络线

Fig.5 The trend of the variables and their envelopes

变量  $\alpha$  与各过程变量之间的相关性, 我们计算这些变量之间的皮尔逊相关系数, 并将结果展示如图 6. 图中对应元素的值表示两变量之间的相关系数, 如第二行第一列的值表示过程变量  $x^1$  与虚拟变量  $\alpha$  之间的相关系数为 0.912 8. 图 6 结果表明, 质量变量  $y_t$  与变量  $x^1 \sim x^4$  的相关系数均高于其与  $x^5 \sim x^7$  的相关系数, 这一规律在 QRV 也同样存在. 该结果与仿真模型 (30) 的实验设置相符. 更重要的是, 质量变量  $y_t$  与质量关联虚拟变量  $\alpha$  的相关系数高达 0.930 8, 这一结果充分表明两者之间具有较强的相关性.

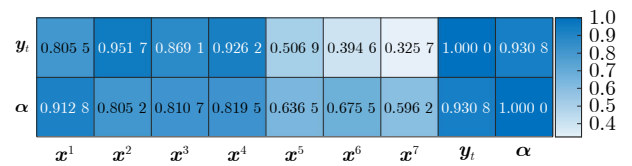


图 6 各变量之间的相关系数

Fig.6 Correlation coefficients between variables

4.1.3 故障检测实验

在模型 (30) 基础上注入故障如下:

$$x_t = x_t^* + \xi \tag{31}$$

故障 1  $\xi = [3 \ 2 \ 4 \ 3 \ 0 \ 0 \ 0]$ ;

故障 2  $\xi = [0 \ 0 \ 0 \ 0 \ 1 \ 2 \ 1]$ .

式 (31) 中  $x_t^*$  表示正常样本;  $\xi$  表示故障分量.

显然, 故障 1 为质量相关故障, 而故障 2 为质量无关故障. 每种故障生成 1 000 个测试样本: 前 300 个为正常样本, 后 700 个为故障样本. 用相同的实验数据, 分别采用 MI-KPCA, MKICR, VIP-DCPLS, OMDPLS, QRVICA-without-EWMA 和 QRVICA 进行故障检测, 结果如图 7 和图 8 所示, 图中  $\hat{T}^2$ ,  $\hat{I}^2$ ,  $d_r$ ,  $T^2$  和  $\varphi_r$  表示质量相关统计量, 而  $\tilde{T}^2$ ,  $\tilde{I}^2$ ,  $d_u$ ,  $Q$  和  $\varphi_u$  表示质量无关统计量.

由图 7 可知, QRVICA 在故障注入之前 (即采样点 1~300), 样本的检测指标均低于控制限, 而故障注入后 (即采样点 301~1 000), 样本的检测指标均高于控制限. 而其他方法均存在不同程度的误报和漏报. 据统计, MI-KPCA, MKICR, VIP-DCPLS, OMDPLS 及 QRVICA-without-EWMA 对故障 1 的 FDR 分别为 84.71%, 76.86%, 75.50%, 77.51% 和 90.00%.

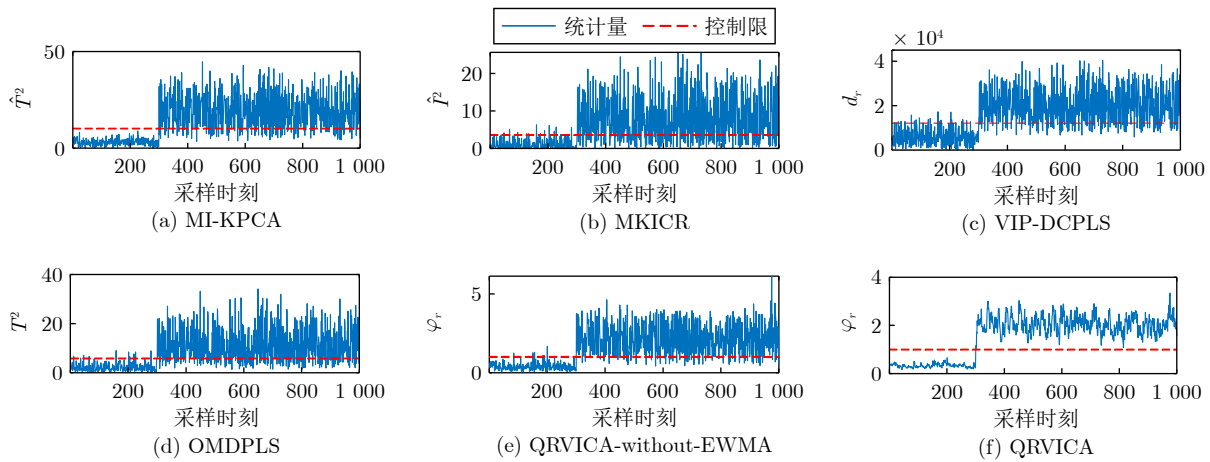


图 7 故障 1 的故障检测结果  
Fig.7 Fault detection results of fault 1

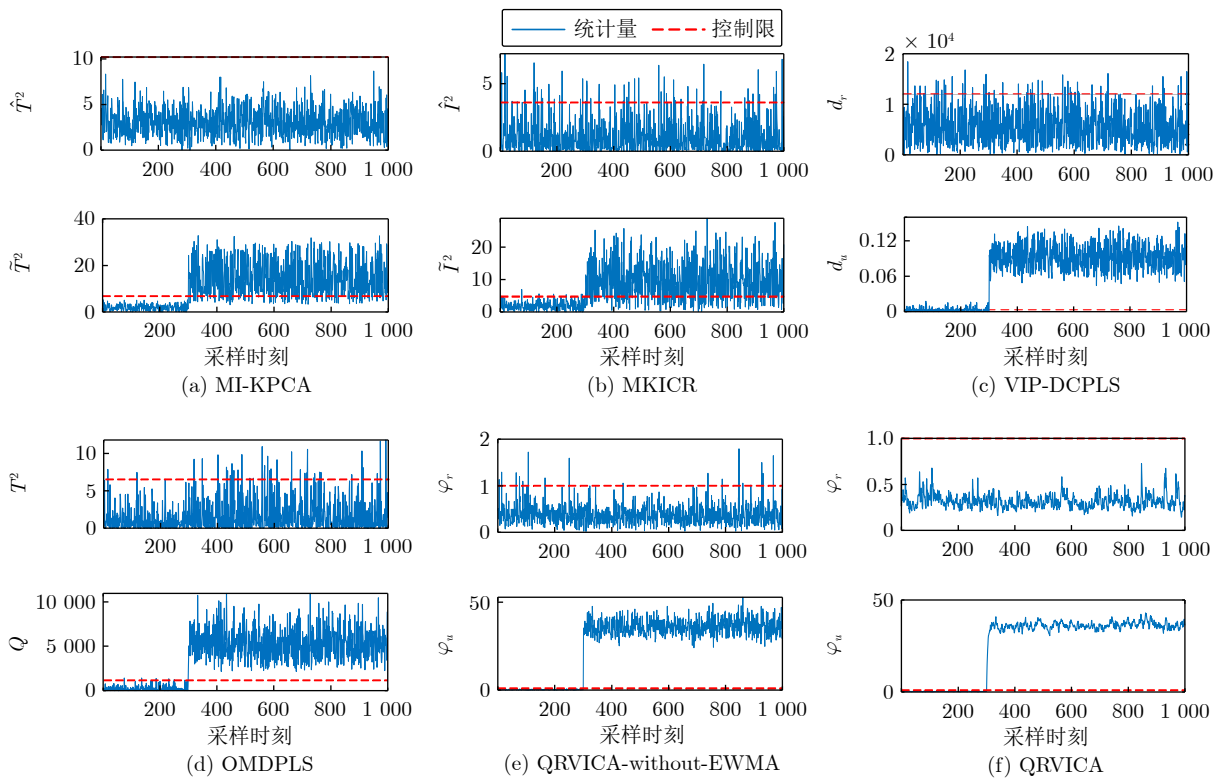


图 8 故障 2 的故障检测结果  
Fig.8 Fault detection results of fault 2

由图 8 可知, MI-KPCA 和 MKICR 中部分故障样本的质量无关检测指标低于控制限, 即存在漏报; 而 MKICR, VIP-DCPLS, OMDPLS 及 QR-VICA-without-EWMA 中一些正常样本的质量相关统计量高于控制限, 导致误报. 相比之下, QRVICA 方法对质量相关和质量无关故障均保持零误报率和 100% 的 FDR. 结合图 4 可知, 尽管 MI-KPCA 和 QRVICA 的变量划分结果一致, 但基于 ICA 的 QRVICA 的 FDR 高于基于 PCA 的 MI-KPCA 的 FDR. 作为另一种 ICA 的变体, 基于空间分解的 MKICR 的 FAR 高于基于变量划分的 QRVICA 的 FAR, 而 FDR 低于 QRVICA 的 FDR. 对比图 7(e) 和图 7(f) 以及图 8(e) 和图 8(f) 可知, 引入 EWMA 后质量相关和质量无关故障的检测性能均有显著提升.

综上, QRVICA 具有优越检测性能的原因主要包括以下几点: 1) QRVICA 作为 ICA 的变体, 能够充分利用数据的高阶统计信息; 2) QRVICA 是基于变量划分思想的质量相关故障检测方法, 其从根本上 (即从建模变量的选择开始) 实现质量相关与质量无关信息的有效分离; 3) EWMA 技术通过融入历史样本信息, 显著提升故障检测性能.

## 4.2 TEP 实验

TEP 是 Eastman 公司根据实际化工反应过程开发的仿真平台, 其被广泛用于测试复杂工业过程的故障检测和诊断模型<sup>[27]</sup>. TEP 包含 41 个测量变量 XMEAS (1 ~ 41) 和 11 个操纵变量 XMV (1 ~ 11). 在本实验中, 选择 XMEAS (1 ~ 22) 和 XMV (1 ~ 11) 组成 33 个过程变量, XMEAS (35) (主要产物 G 的浓度) 作为质量变量  $y$ . 训练集包含 500 个正常样本, 每种故障的测试样本中前 160 个为正常样本, 后 800 个为故障样本.

### 4.2.1 质量相关变量划分实验

按照第 2.2 节所述质量相关变量划分方法将过程变量划分为质量相关变量组  $X_r$  和质量无关变量组  $X_u$ . 设置窗口宽度  $M = 200$ , 划分结果如图 9 所示, 图中用橙色标记质量相关变量, 用蓝色标记质量无关变量. 由图 9 可知, 质量相关变量由测量变量 XMEAS (1, 3, 7, 10, 11, 13, 16, 18, 19, 20) 和操纵变量 XMV (2, 3, 5, 6, 9) 组成. 文献 [17] 计算过程变量与  $y$  之间的 MI 值, 据此得出质量相关变量由测量变量 XMEAS (1, 7, 10, 11, 13, 16, 18, 19, 20) 和操纵变量 XMV (3, 5, 6, 9) 组成. 两种方法中 31 个变量的划分结果一致, 这表明利用 QRV 进行变量划分与直接使用  $y$  的划分结果高度吻合.

TEP 共包括 15 种已知故障模式, 分别记作

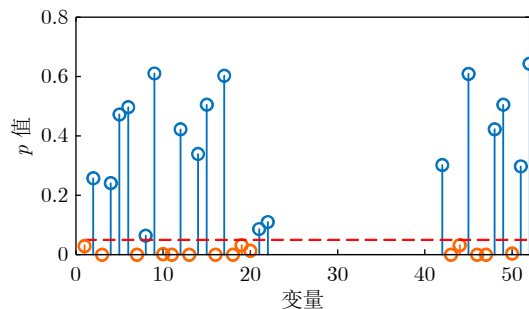


图 9 TEP 过程变量划分结果

Fig.9 Division results of process variable in TEP

IDV (1) ~ IDV (15). 按照文献 [6] 的分类准则, IDV (1, 2, 5 ~ 8, 10, 12, 13) 为质量相关故障; IDV (3, 4, 9, 11, 14, 15) 为质量无关故障, 其中 IDV (3, 9, 15) 中由于变量的波动幅值很小, 大多数常规方法难以检测到故障, 本文不做讨论.

为进一步验证变量划分结果的准确性, 本文以 IDV (5) 为例, 分别用  $X_r$  和  $X_u$  对质量变量  $y$  进行最小二乘回归, 回归结果如图 10, 图中蓝色线条  $y$  表示产物 G 浓度的真实变化趋势, 红色线条  $\hat{y}_r$  和绿色线条  $\hat{y}_u$  分别表示用  $X_r$  和  $X_u$  进行回归的预测值. 由图可知,  $\hat{y}_r$  紧密跟随  $y$  的变化趋势, 这表明  $X_r$  可有效预测  $y$ ; 相反,  $\hat{y}_u$  与  $y$  之间存在显著偏离, 这表明  $X_u$  无法准确预测  $y$  的变化, 即  $X_u$  与  $y$  的相关性较小. 这一结果间接验证了变量划分结果的准确性.

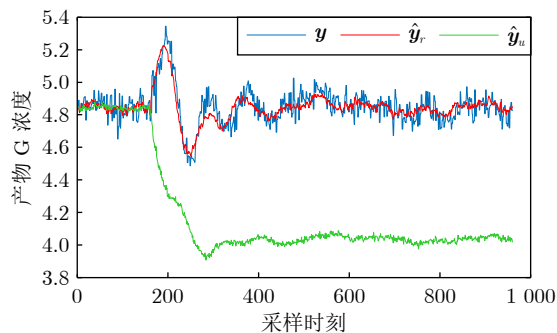


图 10 IDV (5) 中不同变量组对质量变量的回归结果  
Fig.10 Regression results of the quality variable for different variable groups in IDV (5)

### 4.2.2 故障检测实验

#### 1) 质量相关故障检测结果分析

利用 MI-KPCA, MKICR, VIP-DCPLS, OMDPLS, QRVICA-without-EWMA 和 QRVICA 方法, 我们对质量相关故障进行故障检测, 结果如表 1 所示. 对于质量相关故障, 一般更关注质量相关检测指标, 因此表 1 中 FAR 和 FDR 都是对质量相关检

表 1 质量相关故障的 FARs 和 FDRs (%)  
Table 1 FARs and FDRs of the quality-related faults (%)

算法 故障编号	MI-KPCA		MKICR		VIP-DCPLS		OMDPLS		QRVICA-without-EWMA		QRVICA	
	FAR ( $\hat{T}^2$ )	FDR ( $\hat{T}^2$ )	FAR ( $\hat{I}^2$ )	FDR ( $\hat{I}^2$ )	FAR ( $d_r$ )	FDR ( $d_r$ )	FAR ( $T^2$ )	FDR ( $T^2$ )	FAR ( $\varphi_r$ )	FDR ( $\varphi_r$ )	FAR ( $\varphi_r$ )	FDR ( $\varphi_r$ )
1	0.30	99.62	0.63	69.13	<b>0.00</b>	<b>99.75</b>	2.50	99.63	<b>0.00</b>	99.75	<b>0.00</b>	<b>99.75</b>
2	<b>0.00</b>	<b>98.50</b>	0.63	96.88	<b>0.00</b>	96.50	0.63	97.75	<b>0.00</b>	98.25	<b>0.00</b>	<b>98.50</b>
5	<b>0.00</b>	24.12	0.63	20.62	0.63	24.00	1.25	19.25	<b>0.00</b>	24.13	<b>0.00</b>	25.12
6	<b>0.00</b>	99.75	<b>0.00</b>	<b>100.00</b>	<b>0.00</b>	<b>100.00</b>	3.13	98.75	<b>0.00</b>	<b>100.00</b>	<b>0.00</b>	<b>100.00</b>
7	<b>0.00</b>	<b>40.75</b>	0.63	35.13	<b>0.00</b>	40.88	1.87	64.50	<b>0.00</b>	37.00	<b>0.00</b>	37.62
8	2.40	<b>97.62</b>	3.75	76.00	1.25	<b>97.75</b>	1.25	88.13	<b>0.00</b>	97.63	<b>0.00</b>	97.50
10	<b>0.00</b>	79.87	<b>0.00</b>	63.38	<b>0.00</b>	55.63	0.63	54.00	0.63	80.63	<b>0.00</b>	<b>84.88</b>
12	1.25	98.88	21.88	74.13	<b>0.00</b>	98.75	1.25	83.88	<b>0.00</b>	99.25	<b>0.00</b>	<b>99.62</b>
13	<b>0.00</b>	94.63	1.25	85.38	<b>0.00</b>	<b>95.25</b>	0.63	93.75	<b>0.00</b>	93.75	<b>0.00</b>	94.75
平均	0.44	81.53	3.27	68.96	0.21	78.72	1.46	77.74	0.07	81.15	<b>0.00</b>	<b>81.97</b>

测指标的统计结果. 此外, 表 1 中所有方法对 IDV (5) 和 IDV (7) 的检测率均较低, 这是因为这两种故障属于短期故障. 其特点为故障发生后, 质量变量会出现短暂的波动; 随后, 在容错控制器的负反馈作用下, 质量变量会逐渐恢复到正常状态<sup>[14, 17]</sup>. 因此, 对于 IDV (5) 和 IDV (7), 并非检测率越高越好, 而是需要在保证检测准确性的同时避免过度检测.

表 1 中, 除 QRVICA 外, 所有其他方法均存在不同程度的误报. 而 QRVICA 不仅在所有故障中实现零误报, 且在大多数故障中的 FDR 均达到最高水平. 具体而言, 相较于 MI-KPCA, MKICR, VIP-DCPLS, OMDPLS 及 QRVICA-without-EWMA, QRVICA 的平均 FDR 分别提高 0.44, 13.01, 3.25, 4.23 和 0.82 个百分点. 这一优势在 IDV (10) 中尤为明显, 如图 11 所示. 为清晰展示检测结果, 图中

对关键区域进行局部放大处理 (如绿色虚线框内所示). 由图 11 可知, 采用 QRVICA 方法时仅有少量故障样本的检测指标低于控制限. 而采用其他方法时较多故障样本的检测指标低于控制限, 导致这些方法对 IDV (10) 的故障检测率较低. 对比图 11(e) 和图 11(f) 可知, 引入 EWMA 可有效提升故障检测性能, FAR 从 0.63% 降低至 0.00%, 而 FDR 从 80.63% 提高至 84.88%.

2) 质量无关故障检测结果分析

对于质量无关故障, 我们分别利用 MI-KPCA, MKICR, VIP-DCPLS, OMDPLS, QRVICA-without-EWMA 和 QRVICA 方法进行故障检测, 结果如表 2 所示, 其中粗体表示最佳性能. 分析表 2 可知, 在处理质量无关故障时, 相较于基于空间分解的 MKICR 和 OMDPLS 方法, 基于变量划分的

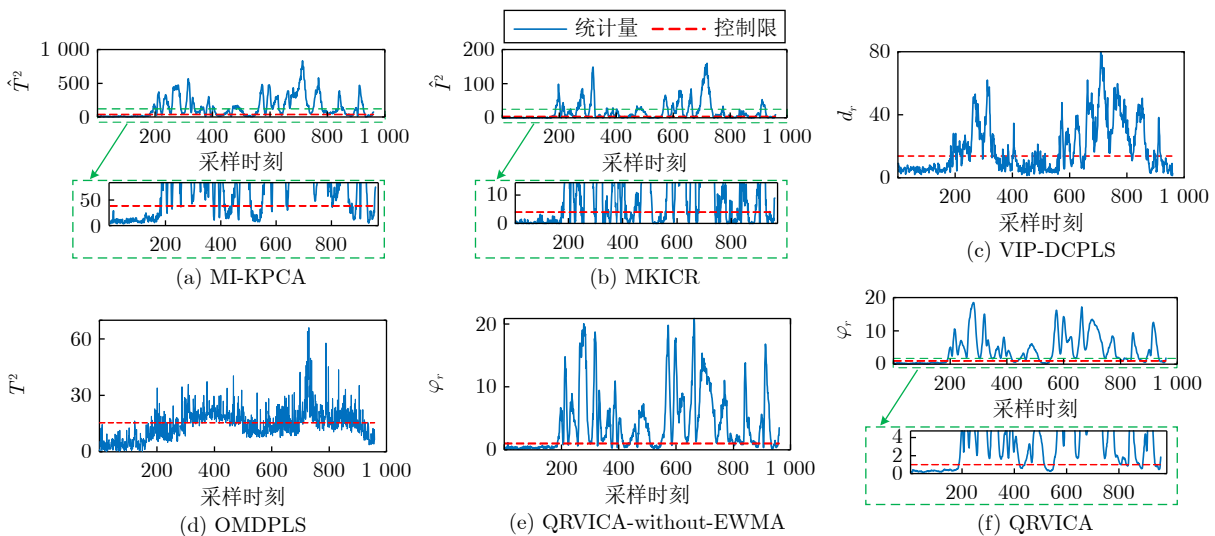


图 11 IDV (10) 的故障检测结果

Fig. 11 Fault detection results of the IDV (10)

表 2 质量无关故障的 FARs 和 FDRs (%)  
Table 2 FARs and FDRs of the quality-unrelated faults (%)

算法 故障编号	MI-KPCA		MKICR		VIP-DCPLS		OMDPLS		QRVICA-without-EWMA		QRVICA	
	FAR ( $\hat{T}^2$ )	FDR ( $\hat{T}^2$ )	FAR ( $\hat{I}^2$ )	FDR ( $\hat{I}^2$ )	FAR ( $d_r$ )	FDR ( $d_u$ )	FAR ( $T^2$ )	FDR ( $Q$ )	FAR ( $\varphi_r$ )	FDR ( $\varphi_u$ )	FAR ( $\varphi_r$ )	FDR ( $\varphi_u$ )
4	0.10	<b>100.00</b>	3.75	<b>100.00</b>	1.25	99.00	13.00	<b>100.00</b>	0.10	<b>100.00</b>	<b>0.00</b>	<b>100.00</b>
11	2.40	69.87	12.00	83.63	2.62	66.50	17.88	79.50	0.31	73.88	<b>0.00</b>	<b>83.75</b>
14	<b>0.00</b>	<b>100.00</b>	17.88	<b>100.00</b>	0.37	<b>100.00</b>	99.38	99.88	0.21	<b>100.00</b>	<b>0.00</b>	<b>100.00</b>
平均	0.83	89.96	11.21	94.54	1.41	88.50	43.42	93.13	0.21	91.29	<b>0.00</b>	<b>94.58</b>

MI-KPCA, VIP-DCPLS 和 QRVICA 方法表现出更低的 FAR. 这主要是因为基于空间分解的方法首先建立过程变量和质量变量之间的回归关系, 然后通过对回归矩阵进行奇异值分解 (Singular value decomposition, SVD), 将空间划分为质量相关子空间和质量无关子空间. 这类方法涉及的环节较多, 包括回归关系的建立、SVD 分解、故障检测策略等. 任一环节的偏差都可能影响故障检测的精度. 而基于变量划分的方法从建模时便实现质量相关与质量无关信息的有效分离, 只要变量划分足够准确, 就能够有效降低误报率. 所提 QRVICA 方法在对 IDV (4), IDV (11) 和 IDV (14) 的检测中零误报, 同时具有最高的故障检测率, 其平均 FDR 相较其他五种方法分别提高 4.62, 0.04, 6.08, 1.45 和 3.29

个百分点.

为更直观地展示检测效果, 图 12 展示 6 种方法对 IDV (11) 的检测结果. 从图中可以看出, 除 QRVICA 外其他方法中部分样本的质量相关检测指标超过控制限, 导致对质量无关故障的误报, 从而可能引发不必要的停机. 而所提 QRVICA 方法的质量相关检测指标  $\varphi_r$  显著低于控制限, 且大多数故障样本的质量无关检测指标  $\varphi_u$  则高于控制限. 实验结果验证了 QRVICA 在故障检测中的优越性.

## 5 结束语

针对工业过程中质量变量难以直接测量的情况, 本文提出一种基于 QRV 的质量相关变量划分方法. 该方法通过构造一个 QRV 作为中间量实现

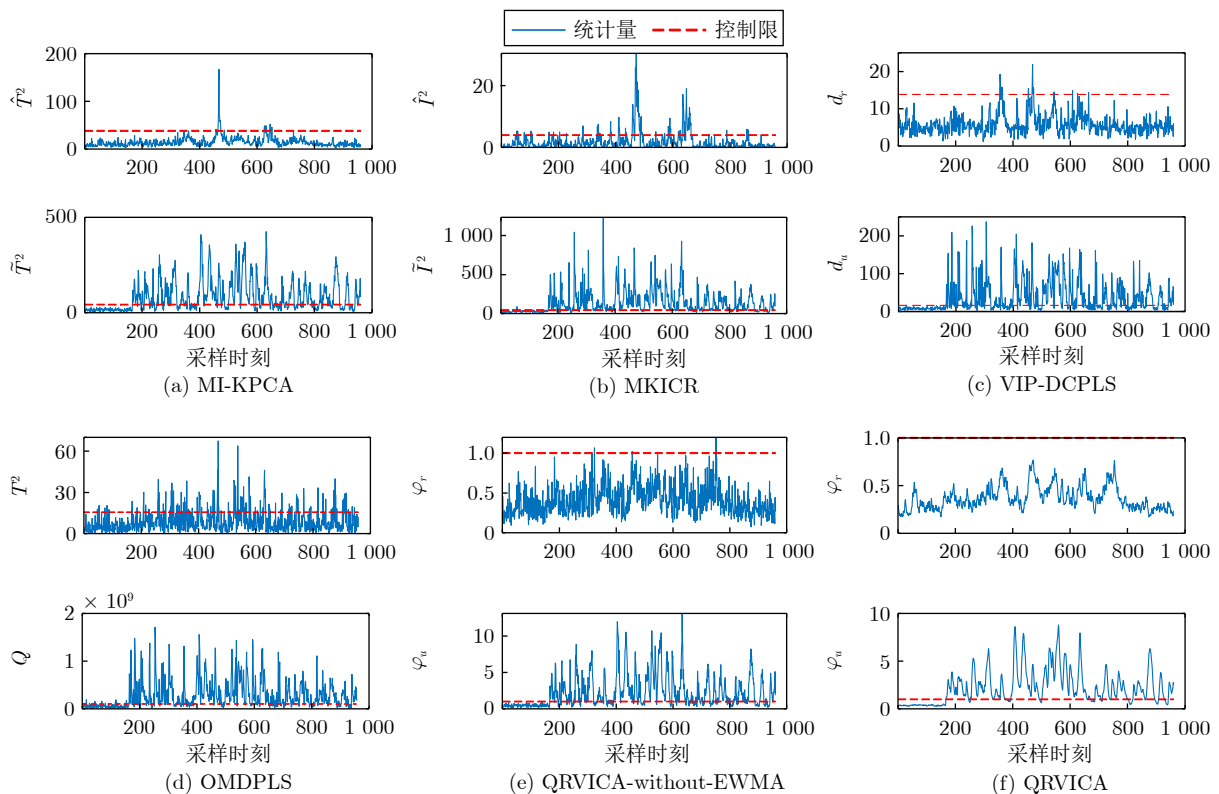


图 12 IDV (11) 的故障检测结果

Fig.12 Fault detection results of the IDV (11)

过程变量的划分,能有效克服传统方法对质量变量的依赖.所提方法特别适用于那些质量变量难以直接测量情况下的质量相关故障检测.此外,该方法也适用于存在多个质量变量的场景,避免传统方法中需逐一评估每对过程变量和质量变量间关系的繁琐过程,进而简化计算流程.将所提质量相关变量划分方法应用到故障检测,实验结果表明,无论是针对质量相关还是质量无关的故障场景,所提 QR-VICA 方法均展现出优越的故障检测性能,不仅能够有效抑制误报率,而且在提升故障检测准确率上具有显著优势,这进一步验证了所提变量划分方法的有效性和可靠性.这一研究为复杂装备或工业过程的质量相关故障检测提供一种新的可行性方案.

值得注意的是,所提方法适用于线性系统,如何将其扩展到非线性系统尚待研究.鉴于非线性系统的复杂性和多样性,未来的研究方向可聚焦于如何将互信息等非线性分析工具有效融入本方法中.

## References

- Dong Jie, Zhang Wei, Peng Kai-Xiang, Ma Liang. A novel method of quality abnormality detection and fault quantitative assessment for industrial processes. *Acta Automatica Sinica*, 2022, **48**(10): 2406–2415  
(董洁, 张伟, 彭开香, 马亮. 一种面向工业过程的质量异常检测与故障量化评估方法. *自动化学报*, 2022, **48**(10): 2406–2415)
- Peng Kai-Xiang, Ma Liang, Zhang Kai. Review of quality-related fault detection and diagnosis techniques in 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)
- Zhou D H, Li G, Qin S J. Total projection to latent structures for process monitoring. *AIChE Journal*, 2010, **56**(1): 168–178
- Ding S X, Yin S, Peng K X, Hao H Y, Shen B. A novel scheme for key performance indicator prediction and diagnosis with application to an industrial hot strip mill. *IEEE Transactions on Industrial Informatics*, 2013, **9**(4): 2239–2247
- Jiao J F, Zhao N, Wang G, Yin S. A nonlinear quality-related fault detection approach based on modified kernel partial least squares. *ISA Transactions*, 2017, **66**: 275–283
- Peng K X, Zhang K, Li G. Quality-related process monitoring based on total kernel PLS model and its industrial application. *Mathematical Problems in Engineering*, 2013, **2013**: Article No. 707953
- Hu C H, Luo J Y, Kong X Y, Xu Z Y. Orthogonal multi-block dynamic PLS for quality-related process monitoring. *IEEE Transactions on Automation Science and Engineering*, 2024, **21**(3): 3421–3434
- Si Y B, Wang Y Q, Zhou D H. Key-performance-indicator-related process monitoring based on improved kernel partial least squares. *IEEE Transactions on Industrial Electronics*, 2021, **68**(3): 2626–2636
- Wang G, Luo H, Peng K X. Quality-related fault detection using linear and nonlinear principal component regression. *Journal of the Franklin Institute*, 2016, **353**(10): 2159–2177
- Chen Q, Wang Y Q. Key-performance-indicator-related state monitoring based on kernel canonical correlation analysis. *Control Engineering Practice*, 2021, **107**: Article No. 104692
- Song Bing, Zheng Cheng-Feng, Shi Hong-Bo, Tao Yang, Tan Shuai. Research on quality-related fault detection method based on VAE-OCCA. *CIESC Journal*, 2023, **74**(4): 1630–1638  
(宋冰, 郑城风, 侍洪波, 陶阳, 谭帅. 基于 VAE-OCCA 的质量相关故障检测方法研究. *化工学报*, 2023, **74**(4): 1630–1638)
- Du W, Fan Y, Zhang Y, Zhang J. Fault diagnosis of non-Gaussian process based on FKICA. *Journal of the Franklin Institute*, 2017, **354**(6): 2573–2590
- Dong J, Wang Y, Peng K. A novel fault detection method based on the extraction of slow features for dynamic nonstationary processes. *IEEE Transactions on Instrumentation and Measurement*, 2022, **71**: Article No. 3500611
- Wang G, Yang J H, Qian Y C, Han J S, Jiao J F. KPCA-CCA-based quality-related fault detection and diagnosis method for nonlinear process monitoring. *IEEE Transactions on Industrial Informatics*, 2023, **19**(5): 6492–6501
- Liu M Z, Kong X Y, Luo J Y, Yang L. Fault detection and diagnosis in a non-Gaussian process with modified kernel independent component regression. *The Canadian Journal of Chemical Engineering*, 2023, **102**(2): 781–802
- Zhang X Y, Ma L, Peng K X, Zhang C F. A quality-related distributed fault detection method for large-scale sequential processes. *Control Engineering Practice*, 2022, **127**: Article No. 105308
- Huang J P, Yan X F. Quality relevant and independent two block monitoring based on mutual information and KPCA. *IEEE Transactions on Industrial Electronics*, 2017, **64**(8): 6518–6527
- Huang J, Yang X, Peng K. Double-layer distributed monitoring based on sequential correlation information for large-scale industrial processes in dynamic and static states. *IEEE Transactions on Industrial Informatics*, 2021, **17**(9): 6419–6428
- Fan H P, Lai X Z, Du S, Yu W K, Lu C D, Wu M. Distributed monitoring with integrated probability PCA and mRMR for drilling processes. *IEEE Transactions on Instrumentation and Measurement*, 2022, **71**: Article No. 3516213
- Yu W K, Zhao C H, Huang B, Xie M. Intrinsic causality embedded concurrent quality and process monitoring strategy. *IEEE Transactions on Industrial Electronics*, 2024, **71**(11): 15111–15121
- Jiao J F, Zhen W T, Zhu W X, Wang G. Quality-related root cause diagnosis based on orthogonal kernel principal component regression and transfer entropy. *IEEE Transactions on Industrial Informatics*, 2021, **17**(9): 6347–6356
- Yang J, Wang J Y, Sha J L, Dai H Q, Liu H B. Quality-related monitoring of distributed process systems using dynamic concurrent partial least squares. *Computers & Industrial Engineering*, 2022, **164**: Article No. 107893
- Yang J, Wang J Y, Ye Q L, Xiong Z X, Zhang F S, Liu H B. A novel fault detection framework integrated with variable importance analysis for quality-related nonlinear process monitoring. *Control Engineering Practice*, 2023, **141**: Article No. 105733
- Guo L, Shi H B, Tan S, Song B, Tao Y. A knowledge-driven spatial-temporal graph neural network for quality-related fault detection. *Process Safety and Environmental Protection*, 2024, **184**: 1512–1524
- Zhu J Z, Shi H B, Song B, Tao Y, Tan S. Convolutional neural network based feature learning for large-scale quality-related process monitoring. *IEEE Transactions on Industrial Informatics*, 2022, **18**(7): 4555–4565
- Lee J M, Yoo C, Lee I B. Statistical process monitoring with independent component analysis. *Journal of Process Control*, 2004, **14**(5): 467–485
- Downs J, Vogel E F. A plant-wide industrial process control problem. *Computers and Chemical Engineering*, 1993, **17**(3): 245–255



刘美枝 山西大同大学讲师, 火箭军工程大学导弹工程学院博士研究生. 2014 年获得西安交通大学硕士学位. 主要研究方向为特征提取, 故障检测与诊断.

E-mail: [lmzdt dx@163.com](mailto:lmzdt dx@163.com)

(LIU Mei-Zhi Lecturer at Shanxi Datong University, Ph.D. candidate at the College of Missile Engineering, Rocket Force University of Engineering. She received her master degree from Xi'an Jiaotong University in 2014. Her research interest covers feature extraction, fault detection and diagnosis.)



孔祥玉 火箭军工程大学导弹工程学院教授. 2005 年获得西安交通大学博士学位. 主要研究方向为自适应信号处理, 特征提取, 故障检测与诊断. 本文通信作者.

E-mail: [xiangyukong01@163.com](mailto:xiangyukong01@163.com)

(KONG Xiang-Yu Professor at the

College of Missile Engineering, Rocket Force University of Engineering. He received his Ph.D. degree from Xi'an Jiaotong University in 2005. His research interest covers adaptive signal processing, feature extraction, fault detection and diagnosis. Corresponding author of this paper.)



胡昌华 火箭军工程大学导弹工程学院教授. 1996 年获得西北工业大学博士学位. 主要研究方向为故障检测与诊断, 寿命预测和容错控制.

E-mail: [hch666@163.com](mailto:hch666@163.com)

(HU Chang-Hua Professor at the College of Missile Engineering, Rocket Force University of Engineering. He received his Ph.D. degree from Northwestern Polytechnical University in 1996. His research interest covers fault detection and diagnosis, life prognosis and fault tolerant control.)