

# 基于并发改进偏最小二乘的质量相关和过程相关的故障诊断

李 强, 孔祥玉<sup>†</sup>, 罗家宇, 解 建

(火箭军工程大学 导弹工程学院, 陕西 西安 710025)

**摘要:** 并发潜结构投影(CPLS)与传统贡献图法是多元统计过程监控中常用的故障检测与诊断方法. 过程监控通常要求监测的时效性与诊断的准确性, 然而, 由于CPLS计算复杂以及传统贡献图诊断结果易受初始贡献较大的变量影响, 因此它们反馈的监控结果可能并不准确. 针对上述问题分别提出一种并发改进偏最小二乘(CMPLS)方法和新的相对贡献图法(NRC). 首先, CMPLS将输入和输出数据同时投影到与过程相关或质量相关的多个子空间, 在相应子空间分别构造适用于各种故障报警的监测指标进行过程监测; 然后, 结合所提出的NRC进行故障识别. 所提方法对过程故障实现全面监测的同时避免了过多的迭代过程, 并消除了过程变量中对检测指标初始贡献较大变量的影响. 最后利用数值仿真和田纳西伊士曼过程验证了所提方法的有效性.

**关键词:** 过程控制; 主元分析; 故障诊断; 偏最小二乘; 安全

**引用格式:** 李强, 孔祥玉, 罗家宇, 等. 基于并发改进偏最小二乘的质量相关和过程相关的故障诊断. 控制理论与应用, 2021, 38(3): 318 – 328

DOI: 10.7641/CTA.2020.00345

## Quality-relevant and process-relevant fault diagnosis with concurrent modified partial least squares

LI Qiang, KONG Xiang-yu<sup>†</sup>, LUO Jia-yu, XIE Jian

(Department of Missile Engineering, Rocket Force University of Engineering, Xi'an Shaanxi 710025, China)

**Abstract:** The concurrent projection to latent structures (CPLS) and conventional contribution plots are often used as fault detection and diagnosis approaches for multivariate statistical process monitoring. Process monitoring usually requires the timeliness of monitoring and the accuracy of diagnosis, but the complex calculation of CPLS and the diagnosis results of conventional contribution plots are susceptible to the variables with large initial contributions, thereby their feedback monitoring results may not be accurate. To solve the above problems, a concurrent modified-PLS (CMPLS) method and a new relative contribution plots (NRC) method are proposed respectively. Originally, CMPLS projects the input and output data to multiple subspaces related to the process or quality at the same time, and respectively constructs monitoring indicators suitable for various fault alarms in the corresponding subspaces to monitor the process; then combined with the proposed NRC to identify the faults. The proposed methods realize comprehensive monitoring of process faults while avoiding too many iterative processes, and eliminates the influence of variables which larger initial contribution in process variables to the detection index. Finally, numerical simulation and Tennessee Eastman process are used to verify the efficiency of the proposed methods.

**Key words:** process control; principal component analysis; fault diagnosis; partial least squares; safety

**Citation:** LI Qiang, KONG Xiangyu, LUO Jiayu, et al. Quality-relevant and process-relevant fault diagnosis with concurrent modified partial least squares. *Control Theory & Applications*, 2021, 38(3): 318 – 328

## 1 引言

随着现代制造业的迅速发展, 工业生产系统的成本与复杂性在快速增长, 相应生产过程的安全性和可靠性也引起了人们的持续关注<sup>[1-2]</sup>. 针对这些工业系

统的需求, 传统的基于机理建模的过程监测与故障诊断方法难以满足现代工业生产. 与之相比, 数据驱动技术可以直接从大量数据中获取状态信息, 不必知道系统的物理模型, 易于实现, 在过程监控领域得到广

收稿日期: 2020-06-11; 录用日期: 2020-09-30.

<sup>†</sup>通信作者. E-mail: xiangyukong01@163.com; Tel.: +86 29-84743331.

本文责任编辑: 周东华.

国家自然科学基金项目(61673387, 61833016), 陕西省自然科学基金项目(2020JM-356)资助.

Supported by the National Natural Science Foundation of China (61673387, 61833016) and the National Natural Science Foundation of Shaanxi Province (2020JM-356).

泛应用<sup>[3-6]</sup>. 多元统计过程监测(multivariate statistical process monitoring, MSPM)是一种有效的数据驱动方法<sup>[7]</sup>, 传统的MSPM投影方法有主成分分析<sup>[8]</sup>(principal component analysis, PCA)和偏最小二乘<sup>[9]</sup>(partial least squares, PLS). Wise等<sup>[10]</sup>和Qin<sup>[11]</sup>对基于PCA和PLS的过程监测和故障诊断方法进行了总结, 这些方法是根据正常操作中收集到的过程和质量数据实现的.

随着分布式控制系统<sup>[12]</sup>和新型测量设备的广泛应用, 过程工业中大量的过程数据和产品质量数据得以记录和储存. 与过程数据测量相比, 质量数据测量频率较低, 通常具有明显的时间延迟. PCA最初是为了对大量相关数据进行降维而提出的, 它能够保留从过程变量中提取的显著变异信息. 因此, 当质量数据难以获得时, PCA被用作监测过程变量的异常变化. 然而在基于PCA的故障检测方法中, 对过程变量进行分解时没有考虑输入与输出变量间的关系, 无法揭示在过程变量中检测到的故障是否与质量输出相关<sup>[13]</sup>. 投入产出相关性可以为过程监控提供更多有用的信息<sup>[14]</sup>, 适用于此的一种方法是基于PLS的故障检测方法. PLS通过提取质量变量与过程变量之间的最大相关性建立了一个输入-输出模型, 可用于监测与质量输出相关的输入子空间, 从而判断输入子空间是否发生故障, 故障是否对质量输出产生影响.

实际上, PLS在通过监测得分进行质量相关故障检测和通过监测输入残差来检测质量无关故障方面存在如下缺陷<sup>[13]</sup>: 1) PLS对可用的输入数据进行了不期望的分解, 利用PLS得分构造的 $T^2$ 统计量中包含了与质量变量正交的信息; 2) PLS并非按照降序提取输入与输出变量间的协方差, 因此输入残差中可能包含较大的变化, 不适合使用 $Q$ 统计量进行监测. 为了解决上述问题, Zhou等<sup>[15]</sup>于2010年提出全潜结构投影(total projection to latent structures, TPLS)方法, 进一步分解了PLS主元空间和残差空间. 与标准PLS方法相比, 该方法大大提高了质量相关故障检测的性能. Qin等<sup>[13]</sup>认为TPLS将输入数据空间过多的分解为4个子空间, 从对质量变量全局监控的角度, 构建了并发潜结构投影(concurrent projection to latent structures, CPLS)模型, 简化了TPLS模型的结构. 但是CPLS存在大量的迭代过程, 计算复杂, 这同样使得基于CPLS的其他拓展(如非线性, 动态等)方法也存在大量迭代过程. 因此, 急需解决CPLS的这一不足. Yin等<sup>[16]</sup>提出了改进潜结构投影(modified-PLS, MPLS)模型, 实现了过程变量空间的正交分解, 弥补了PLS不足的同时采用SVD分解避免了大量的迭代过程. Zhang等<sup>[17]</sup>研究指出, MPLS由于 $\mathbf{X}^T \mathbf{X}$ 的广义逆计算, 可能导致 $\mathbf{X}$ 中与 $\mathbf{Y}$ 相关的信息丢失, 因此提出了高效潜结构投影算法(efficient projection to latent structures,

EPLS). 尽管MPLS与EPLS在检测质量相关故障中表现出了较之PLS更好的性能, 但是难以有效地对过程相关故障以及不可预测的输出残差子空间进行区分与监控, 同时MPLS无法对影响输入数据空间并可能会成为输出潜在故障进行区分. 针对上述问题本文提出一种并发改进偏最小二乘(concurrent modified partial least squares, CMPLS)算法和一组过程监测指标. 所提方法根据统计过程的监测结果构造故障检测的控制限, 对质量相关故障和过程相关故障进行完整检测, 同时避免过多的迭代过程.

当检测到故障后, 更重要的是诊断出故障原因. 贡献图<sup>[18-19]</sup>及其改进的方法<sup>[20-21]</sup>在这一问题的解决上被广泛应用, 它表征了各过程变量对故障检测指标的贡献. 贡献图最先由MacGregor等<sup>[22]</sup>提出, 先后运用到各领域的过程控制与判别分析中<sup>[23]</sup>. 虽然贡献图可以有效地诊断出对质量输出影响比较大的故障变量, 但是研究发现, 在无故障时每个过程输入变量对故障检测指标的贡献值并不是均等的. 使用贡献图进行诊断时有可能受到本来贡献较大变量的影响, 使得诊断结果不可解释. 本文通过正常训练样本的各变量向监测指标的贡献对测试样本各变量向监测指标的贡献进行归一化处理, 提出一种新的相对贡献图法(new relative contribution, NRC). 当CMPLS检测到故障后, 利用NRC进行故障识别, 消除初始贡献较大的过程输入变量的影响, 得到准确的诊断结果. 最后通过仿真实验验证所提算法的故障检测性能和故障诊断性能. 本文的创新点体现在以下几方面: 1) 提出了CMPLS算法划分新的投影空间, 弥补了MPLS不足的同时避免大量迭代过程; 2) 对贡献图进行了改进, 提出了一种NRC方法进行故障诊断, 消除了过程变量中对检测指标初始贡献较大的影响; 3) 结合所提CMPLS算法和NRC方法提出了一套完整的故障检测与诊断技术.

本文剩余部分结构如下: 第2节简要介绍基于MPLS的故障检测方法; 第3节给出基于CMPLS的质量相关和过程相关故障检测方法的详细过程, 定义并推导出各故障监测指标的变量贡献, 进行故障诊断; 第4节通过数值仿真案例与田纳西伊士曼过程(tennessee eastman process, TEP), 同MPLS与EPLS方法进行比较分析, 说明所提方法的有效性; 第5节给出本文结论.

## 2 基于MPLS的故障检测方法

给出一组过程输入数据 $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_m] \in \mathbb{R}^{n \times m}$ 和质量输出数据 $\mathbf{Y} = [\mathbf{y}_1 \ \mathbf{y}_2 \ \cdots \ \mathbf{y}_l] \in \mathbb{R}^{n \times l}$ ,  $\mathbf{X}$ 具有 $n$ 个样本点,  $m$ 个变量;  $\mathbf{Y}$ 具有 $n$ 个样本点,  $l$ 个变量. 其中:  $\mathbf{x}_i = [x_{1i} \ x_{2i} \ \cdots \ x_{ni}]^T \in \mathbb{R}^n$ ,  $i = 1, 2, \cdots, m$ ;  $\mathbf{y}_j = [y_{1j} \ y_{2j} \ \cdots \ y_{lj}]^T \in \mathbb{R}^n$ ,  $j = 1, 2, \cdots, l$ . 通常, 为了技术上的便利以及消除变量的量纲

效应,需要对数据进行零均值、单位方差标准化处理,标准化后的数据仍记为 $\mathbf{X}$ ,  $\mathbf{Y}$ .

针对PLS的不足, MPLS对 $\mathbf{X}$ 进行正交分解,将 $\mathbf{X}$ 成功分解为与 $\mathbf{Y}$ 相关的 $\hat{\mathbf{X}}$ 和与 $\mathbf{Y}$ 无关的 $\tilde{\mathbf{X}}$ . MPLS的模型如下:

$$\begin{cases} \mathbf{X} = \hat{\mathbf{X}} + \tilde{\mathbf{X}}, \\ \mathbf{Y} = \hat{\mathbf{Y}} + \tilde{\mathbf{Y}} = \mathbf{X}\mathbf{M} + \tilde{\mathbf{Y}}, \end{cases} \quad (1)$$

其中 $\mathbf{M} = (\mathbf{X}^T \mathbf{X})^\dagger \mathbf{X}^T \mathbf{Y} \in \mathbb{R}^{m \times l}$ 为 $\mathbf{X}$ 与 $\mathbf{Y}$ 的回归系数矩阵,包含 $\mathbf{X}$ 与 $\mathbf{Y}$ 之间的相关关系.  $\mathbf{Y}$ 通过 $\mathbf{M}$ 矩阵分解为 $\hat{\mathbf{Y}}$ 与 $\tilde{\mathbf{Y}}$ ,其中 $\hat{\mathbf{Y}}$ 与 $\mathbf{X}$ 相关, $\tilde{\mathbf{Y}}$ 与 $\mathbf{X}$ 不相关.然后将 $\mathbf{X}$ 向 $\text{span}\{\mathbf{M}\}$ 和 $\text{span}\{\mathbf{M}\}^\perp$ 投影,得到正交的 $\hat{\mathbf{X}}$ 与 $\tilde{\mathbf{X}}$ ,MPLS算法具体步骤如下:

标准化数据矩阵 $\mathbf{X}$ 和 $\mathbf{Y}$ :

**Step 1** 计算 $\mathbf{M}$ ,  $\mathbf{M} = (\mathbf{X}^T \mathbf{X})^\dagger \mathbf{X}^T \mathbf{Y}$ ;

**Step 2** 对 $\mathbf{M}\mathbf{M}^T$ 进行SVD分解得到 $\hat{\Gamma}_\varphi$ 和 $\tilde{\Gamma}_\varphi$ ,

$$\mathbf{M}\mathbf{M}^T = \begin{bmatrix} \hat{\Gamma}_\varphi & \tilde{\Gamma}_\varphi \end{bmatrix} \begin{bmatrix} \Lambda_\varphi & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{\Gamma}_\varphi^T \\ \tilde{\Gamma}_\varphi^T \end{bmatrix};$$

**Step 3** 计算 $\Pi_\varphi = \hat{\Gamma}_\varphi \hat{\Gamma}_\varphi^T$ ,  $\Pi_\varphi^\perp = \tilde{\Gamma}_\varphi \tilde{\Gamma}_\varphi^T$ ;

**Step 4**  $\hat{\mathbf{X}} = \mathbf{X}\Pi_\varphi$ ,  $\tilde{\mathbf{X}} = \mathbf{X}\Pi_\varphi^\perp$ ,  $\hat{\mathbf{Y}} = \mathbf{X}\mathbf{M}$ .

以上算法中 $\hat{\mathbf{X}} = \mathbf{X}\Pi_\varphi$ 与 $\tilde{\mathbf{X}} = \mathbf{X}\Pi_\varphi^\perp$ 可作如下处理:

$\hat{\mathbf{X}} = \mathbf{X}\Pi_\varphi = \mathbf{T}_{\hat{\mathbf{x}}}\hat{\Gamma}_\varphi^T$ ,  $\tilde{\mathbf{X}} = \mathbf{X}\Pi_\varphi^\perp = \mathbf{T}_{\tilde{\mathbf{x}}}\tilde{\Gamma}_\varphi^T$ ,其中: $\mathbf{T}_{\hat{\mathbf{x}}} = \mathbf{X}\hat{\Gamma}_\varphi$ ,  $\mathbf{T}_{\tilde{\mathbf{x}}} = \mathbf{X}\tilde{\Gamma}_\varphi$ .

给定一个新样本 $\mathbf{x}$ ,将其分解为 $\hat{\mathbf{x}}$ 与 $\tilde{\mathbf{x}}$ 两部分,则有

$$\hat{\mathbf{x}}^T \hat{\mathbf{x}} = \mathbf{x}^T \hat{\Gamma}_\varphi \hat{\Gamma}_\varphi^T \hat{\Gamma}_\varphi \hat{\Gamma}_\varphi^T \mathbf{x} = \mathbf{t}_{\hat{\mathbf{x}}}^T \mathbf{t}_{\hat{\mathbf{x}}}, \quad (2)$$

$$\tilde{\mathbf{x}}^T \tilde{\mathbf{x}} = \mathbf{x}^T \tilde{\Gamma}_\varphi \tilde{\Gamma}_\varphi^T \tilde{\Gamma}_\varphi \tilde{\Gamma}_\varphi^T \mathbf{x} = \mathbf{t}_{\tilde{\mathbf{x}}}^T \mathbf{t}_{\tilde{\mathbf{x}}}, \quad (3)$$

其中: $\mathbf{t}_{\hat{\mathbf{x}}} = \hat{\Gamma}_\varphi^T \mathbf{x}$ ,  $\mathbf{t}_{\tilde{\mathbf{x}}} = \tilde{\Gamma}_\varphi^T \mathbf{x}$ .则质量相关统计量为

$$T_{\hat{\mathbf{x}}}^2 = \mathbf{t}_{\hat{\mathbf{x}}}^T \left( \frac{\mathbf{T}_{\hat{\mathbf{x}}}^T \mathbf{T}_{\hat{\mathbf{x}}}}{n-1} \right)^{-1} \mathbf{t}_{\hat{\mathbf{x}}} = \mathbf{x}^T \hat{\Gamma}_\varphi \left( \frac{\hat{\Gamma}_\varphi^T \mathbf{X}^T \mathbf{X} \hat{\Gamma}_\varphi}{n-1} \right) \hat{\Gamma}_\varphi^T \mathbf{x}. \quad (4)$$

质量无关统计量为

$$T_{\tilde{\mathbf{x}}}^2 = \mathbf{t}_{\tilde{\mathbf{x}}}^T \left( \frac{\mathbf{T}_{\tilde{\mathbf{x}}}^T \mathbf{T}_{\tilde{\mathbf{x}}}}{n-1} \right)^{-1} \mathbf{t}_{\tilde{\mathbf{x}}} = \mathbf{x}^T \tilde{\Gamma}_\varphi \left( \frac{\tilde{\Gamma}_\varphi^T \mathbf{X}^T \mathbf{X} \tilde{\Gamma}_\varphi}{n-1} \right) \tilde{\Gamma}_\varphi^T \mathbf{x}. \quad (5)$$

利用 $\chi^2$ 分布<sup>[24]</sup>来计算 $T_{\hat{\mathbf{x}}}^2$ 与 $T_{\tilde{\mathbf{x}}}^2$ 的控制限,相应的计算如下:

$$J_{\text{th}, T_{\hat{\mathbf{x}}}^2} = \hat{g} \chi_{\alpha, \hat{h}}^2, \quad \hat{g} = \frac{\hat{S}}{2\hat{\mu}}, \quad \hat{h} = \frac{2\hat{\mu}^2}{\hat{S}}, \quad (6)$$

$$J_{\text{th}, T_{\tilde{\mathbf{x}}}^2} = \tilde{g} \chi_{\alpha, \tilde{h}}^2, \quad \tilde{g} = \frac{\tilde{S}}{2\tilde{\mu}}, \quad \tilde{h} = \frac{2\tilde{\mu}^2}{\tilde{S}}, \quad (7)$$

其中: $\hat{\mu}$ 和 $\hat{S}$ 是正常训练样本中 $T_{\hat{\mathbf{x}}}^2$ 的均值与方差,  $\tilde{\mu}$ 和 $\tilde{S}$ 是正常训练样本中 $T_{\tilde{\mathbf{x}}}^2$ 的均值与方差.

### 3 基于CMPLS的故障诊断方法

MPLS避免了复杂的迭代过程,展现出了较之PLS更好的性能,但是在MPLS中对 $\mathbf{X}^T \mathbf{X}$ 进行广义逆,可能会使得 $\hat{\mathbf{X}}$ 中掺入与质量输出 $\mathbf{Y}$ 无关的信息,而 $\tilde{\mathbf{X}}$ 可能含有与质量输出 $\mathbf{Y}$ 相关的信息<sup>[17]</sup>.其次,不论监测 $\hat{\mathbf{X}}$ 空间还是 $\tilde{\mathbf{X}}$ 空间,都仅仅监测的是质量输出 $\mathbf{Y}$ 可预测的部分,而忽略了对质量输出 $\mathbf{Y}$ 不可预测部分的监测.基于此,提出了一个能够全面监测质量输出空间和过程输入空间的CMPLS监控方案.

#### 3.1 模型建立

CMPLS主要实现以下目标:

1) 利用MPLS将过程输入空间 $\mathbf{X}$ 分为 $\hat{\mathbf{X}}$ 与 $\tilde{\mathbf{X}}$ ,对 $\hat{\mathbf{X}}$ 进一步投影到输入主子空间(input principal subspace, IPS)与输入主残差子空间(residual subspace of the input principal, IPRS),其中IPS负责监控与质量输出相关的故障.

2) 对 $\tilde{\mathbf{X}}$ 进一步投影到输入残差主子空间(principal subspace of the input residual, IRPS)和输入残差的残差子空间(residual subspace of the input residuals, IRRS),并将IPRS与IRRS合并为输入残差子空间(input residual subspace, IRS),在IRPS中监控影响输入空间并且可能与质量输出相关的潜在故障,在IRS中监控与质量输出无关的故障,如图1所示.

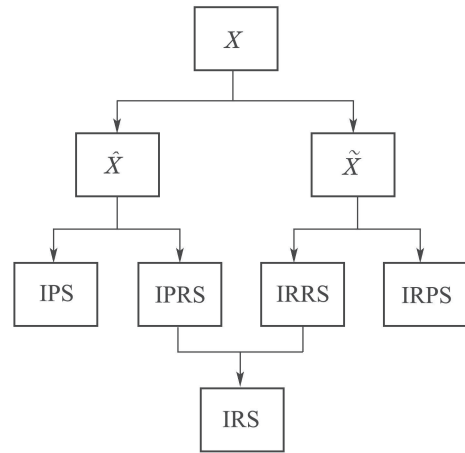


图1 CMPLS对过程输入变量空间的分解图

Fig. 1 Decomposition diagram of CMPLS for process input variable space

3) 将不可预测的输出变化进一步投影到输出主子空间(output principal subspace, OPS)和输出残差子空间(output residual subspace, ORS),以监测与不可预测的质量输出相关的故障.

CMPLS算法的具体过程如下:

1) 将原始数据进行标准化,得到 $\mathbf{X}$ 和 $\mathbf{Y}$ ,利用式(1)对 $\mathbf{X}$ 和 $\mathbf{Y}$ 进行MPLS处理,得到 $\mathbf{T}_{\hat{\mathbf{x}}}$ ,  $\hat{\Gamma}_\varphi$ ,则 $\hat{\mathbf{X}} = \mathbf{T}_{\hat{\mathbf{x}}}\hat{\Gamma}_\varphi^T$ ,  $\tilde{\mathbf{X}} = \mathbf{X} - \hat{\mathbf{X}}$ .

2) 得到不可预测的输出 $\mathbf{Y}_c = \mathbf{Y} - \mathbf{X}\mathbf{M}$ ,并对 $\mathbf{Y}_c$

进行 $l_y$ 个主元的PCA处理:

$$Y_c = T_y P_y^T + \tilde{Y}_c, \quad (8)$$

其中:  $T_y$ 为输出主元得分,  $\tilde{Y}_c$ 为输出残差.

3) 对 $\hat{X} = T_{\hat{x}} \hat{\Gamma}_{\varphi}^T$ 进行 $l_p$ 个主元的PCA处理:

$$\hat{X} = X_p + \tilde{X}_p = T_p P_p^T + \tilde{X}_p, \quad (9)$$

其中:  $T_p$ 为输入主元得分,  $\tilde{X}_p$ 为输入主残差.

4) 对 $\tilde{X}$ 进行 $l_r$ 个主元的PCA处理:

$$\tilde{X} = X_r + \tilde{X}_r = T_r P_r^T + \tilde{X}_r, \quad (10)$$

由此得到输入残差主元得分 $T_r$ 和输入残差的残差 $\tilde{X}_r$ .

5) 将输入主残差 $\tilde{X}_p$ 和输入残差的残差 $\tilde{X}_r$ 合并为输入残差 $\tilde{X}_z$ .

$$\tilde{X}_z = \tilde{X}_p + \tilde{X}_r. \quad (11)$$

通过投影变换得到CMPLS模型如下:

$$\begin{cases} X = T_p P_p^T + T_r P_r^T + \tilde{X}_z, \\ Y = X M + T_y P_y^T + \tilde{Y}_c, \end{cases} \quad (12)$$

其中:  $l_y$ ,  $l_p$  和  $l_r$  分别由累计方差贡献率(cumulative percent variance, CPV)<sup>[25]</sup>确定,  $P_p \in \mathbb{R}^{m \times l_p}$ ,  $P_r \in \mathbb{R}^{m \times l_r}$ ,  $P_y \in \mathbb{R}^{l \times l_y}$ ;  $T_p$ 表示过程输入 $X$ 中与质量输出 $Y$ 的可预测部分 $\hat{Y}$ 相关的信息;  $T_r$ 表示过程输入 $X$ 中与质量输出 $Y$ 相关的潜在故障存在的空间;  $T_y$ 表示质量输出 $Y$ 中无法被过程输入 $X$ 预测的信息;  $\tilde{X}_z$ 表示与过程输入 $X$ 相关, 与质量输出 $Y$ 无关的故障存在空间. 令 $x^T$ ,  $\hat{x}^T$ ,  $\tilde{x}_p^T$ ,  $\tilde{x}^T$ ,  $\tilde{x}_r^T$ ,  $\tilde{x}_z^T$ ,  $y^T$ ,  $y_c^T$ ,  $\tilde{y}^T$ ,  $t_{\hat{x}}^T$ ,  $t_p^T$ ,  $t_r^T$ ,  $t_y^T$ 分别为 $X$ ,  $\hat{X}$ ,  $\tilde{X}_p$ ,  $\tilde{X}$ ,  $\tilde{X}_r$ ,  $\tilde{X}_z$ ,  $Y$ ,  $Y_c$ ,  $\tilde{Y}_c$ ,  $T_{\hat{x}}$ ,  $T_p$ ,  $T_r$ ,  $T_y$ 对应的行向量, 则存在如下关系:

$$\hat{x}^T = t_{\hat{x}}^T \hat{\Gamma}_{\varphi}^T, \quad (13)$$

$$y_c^T = y^T - x^T M, \quad (14)$$

$$\tilde{x}_p^T = \hat{x}^T - t_p^T P_p^T = x^T \hat{\Gamma}_{\varphi} \hat{\Gamma}_{\varphi}^T (I - P_p P_p^T), \quad (15)$$

$$\tilde{x}^T = x^T - t_{\hat{x}}^T \hat{\Gamma}_{\varphi}^T = x^T (I - \hat{\Gamma}_{\varphi} \hat{\Gamma}_{\varphi}^T), \quad (16)$$

$$\tilde{x}_r^T = \tilde{x}^T - t_r^T P_r^T = x^T (I - \hat{\Gamma}_{\varphi} \hat{\Gamma}_{\varphi}^T) (I - P_r P_r^T), \quad (17)$$

$$\tilde{y}^T = y_c^T - t_y^T P_y^T = y_c^T (I - P_y P_y^T), \quad (18)$$

$$\begin{aligned} \tilde{x}_z^T &= \tilde{x}_p^T + \tilde{x}_r^T = \\ &x^T [\hat{\Gamma}_{\varphi} \hat{\Gamma}_{\varphi}^T (I - P_p P_p^T) + (I - \hat{\Gamma}_{\varphi} \hat{\Gamma}_{\varphi}^T) (I - P_r P_r^T)], \end{aligned} \quad (19)$$

其中:

$$t_{\hat{x}} = \hat{\Gamma}_{\varphi}^T x, \quad (20)$$

$$t_p = P_p^T \hat{x}, \quad (21)$$

$$t_r = P_r^T \tilde{x}, \quad (22)$$

$$t_y = P_y^T y_c. \quad (23)$$

式(15)–(23)给出了所有与质量输出相关或过程输

入相关的主元空间变化信息和残差空间变化信息.

### 3.2 故障检测

根据第3.1节中的CMPLS模型, 可以给出故障检测统计指标. 式(21)反应了与质量输出相关的输入数据得分; 式(19)和式(22)分别为与过程输入相关的变化; 式(22)可能含有一些潜在的影响质量输出的变化信息, 即便某些变化可能不会影响质量输出, 但它们可能会导致过程操作中的后续性能损失, 并不希望让这些变化不受控制<sup>[13]</sup>. 另外, 式(23)和式(18)是必须监测的不可预测的质量输出变化和残差, 这是多变量质量监测的一项基本任务<sup>[26]</sup>.

给出一个新样本 $x_{new}$ , 由第3.1节中式(13)–(23)计算出所有变化信息. IPS与IRPS中采用 $T^2$ 统计量进行监控, 如式(24)与式(25):

$$T_{x_p}^2 = t_{p,new}^T \left( \frac{T_p^T T_p}{n-1} \right)^{-1} t_{p,new}, \quad (24)$$

其中:  $t_{p,new} = P_p^T \hat{x}_{new}$ ,  $\hat{x}_{new}^T = t_{\hat{x},new}^T \hat{\Gamma}_{\varphi}^T$ ,  $t_{\hat{x},new} = \hat{\Gamma}_{\varphi}^T x_{new}$ ;

$$T_{x_r}^2 = t_{r,new}^T \left( \frac{T_r^T T_r}{n-1} \right)^{-1} t_{r,new}, \quad (25)$$

其中:  $t_{r,new} = P_r^T \tilde{x}_{new}$ ,  $\tilde{x}_{new}^T = x_{new}^T - t_{\hat{x},new}^T \hat{\Gamma}_{\varphi}^T = x_{new}^T (I - \hat{\Gamma}_{\varphi} \hat{\Gamma}_{\varphi}^T)$ ,  $t_{\hat{x},new} = \hat{\Gamma}_{\varphi}^T x_{new}$ .

IRS中采用 $Q$ 统计量进行监控, 如式(26):

$$Q_{\tilde{x}_z} = \|\tilde{x}_{z,new}\|^2, \quad (26)$$

其中:

$$\begin{aligned} \tilde{x}_{z,new} &= \tilde{x}_{p,new}^T + \tilde{x}_{r,new}^T = \\ &x_{new}^T [\hat{\Gamma}_{\varphi} \hat{\Gamma}_{\varphi}^T (I - P_p P_p^T) + (I - \hat{\Gamma}_{\varphi} \hat{\Gamma}_{\varphi}^T) (I - P_r P_r^T)], \\ \tilde{x}_{p,new}^T &= \hat{x}_{new}^T - t_{p,new}^T P_p^T, \\ \tilde{x}_{r,new}^T &= \tilde{x}_{new}^T - t_{r,new}^T P_r^T. \end{aligned}$$

OPS与ORS中分别采用 $T^2$ 统计量与 $Q$ 统计量进行监测, 如式(27)与式(28):

$$T_y^2 = t_{y,new}^T \left( \frac{T_y^T T_y}{n-1} \right)^{-1} t_{y,new}, \quad (27)$$

其中:  $t_{y,new} = P_y^T y_{c,new}$ ,  $y_{c,new}^T = y_{new}^T - x_{new}^T M$ ;

$$Q_y = \|\tilde{y}_{new}\|^2, \quad (28)$$

其中 $\tilde{y}_{new}^T = y_{c,new}^T - t_{y,new}^T P_y^T = y_{c,new}^T (I - P_y P_y^T)$ .

式(24)–(28)中各统计量的阈值 $J_{th, T_{x_p}^2}$ ,  $J_{th, T_{x_r}^2}$ ,  $J_{th, Q_{\tilde{x}_z}}$ ,  $J_{th, T_y^2}$ ,  $J_{th, Q_y}$ 分别由式(6)或式(7)的形式在训练数据中计算的各统计量中确定.

故障检测策略总结如下:

1) 如果 $T_{x_p}^2 > J_{th, T_{x_p}^2}$ , 则基于过程输入数据 $x_{new}$ 检测到与质量输出相关的故障;

2) 如果 $T_{x_p}^2 < J_{th, T_{x_p}^2}$ 且 $Q_{\tilde{x}_z} > J_{th, Q_{\tilde{x}_z}}$ , 则基于过程输入数据 $x_{new}$ 检测到与质量输出无关, 但与过程输

入相关的故障;

3) 如果  $T_{x_p}^2 < J_{th, T_{x_p}^2}$  且  $T_{x_r}^2 > J_{th, T_{x_r}^2}$ , 则基于过程输入数据  $\mathbf{x}_{new}$  检测到一个可能与质量输出相关的潜在故障, 一般不直接影响质量输出, 而体现在无法预测的质量输出变化中, 通常被认为与质量输出无关的故障;

4) 当质量输出变量可获得时, 如果  $T_y^2 > J_{th, T_y^2}$  或  $Q_y > J_{th, Q_y}$ , 其他部分统计量大于控制限, 则检测到与过程输入及不可预测的质量输出相关的故障; 若  $T_y^2 > J_{th, T_y^2}$  或  $Q_y > J_{th, Q_y}$ , 其他统计量均小于控制限, 则检测到与过程输入无关, 但与不可预测的质量输出相关的故障。

对于上述4)中的检测策略, 由于  $T_y^2$  和  $Q_y$  均反映的是与质量输出相关的不可预测的信息变化, 因此可以更加简单的使用如式(29)所示的综合统计指标进行监测:

$$F_y = \frac{T_y^2}{J_{th, T_y^2}} + \frac{Q_y}{J_{th, Q_y}}. \quad (29)$$

综合统计指标  $F_y$  的控制限  $J_{th, F_y}$  由式(6)或式(7)的形式在训练数据中计算的综合统计指标进行确定. 为了更加直观, 以相对统计量代替绝对统计量进行过程监测, 相对统计量如式(30)所示, 则相对统计量的控制限均为1:

$$\text{相对统计量} = \frac{\text{统计量}}{\text{控制限}}. \quad (30)$$

### 3.3 故障诊断

当统计量检测到故障时, 需要进一步分析, 以诊断出哪些变量可能会受该故障影响. 贡献图是基于数据驱动故障诊断中最流行的一种方法<sup>[27-28]</sup>. 传统的贡献图方法中部分过程变量在系统正常运行时对统计量的贡献就比较大, 这会对故障诊断产生干扰, 无法准确定位故障的变量. 因此, 本节提出了一种新的相对贡献图方法进行故障诊断.

基于  $Q_{\tilde{x}_z}$  统计量的贡献定义如下:

$$C(Q_{\tilde{x}_z}, k) = \tilde{x}_{z, new, k}^2, \quad (31)$$

其中:  $C(Q_{\tilde{x}_z}, k)$  表示第  $k$  个变量对  $Q_{\tilde{x}_z}$  统计量的贡献,  $\tilde{x}_{z, new, k}$  表示  $\tilde{\mathbf{x}}_{z, new}$  的第  $k$  个元素.

基于  $T_{x_p}^2$  统计量的贡献定义如下<sup>[29]</sup>:

$$C(T_{x_p}^2, k) = \left\| \left( \frac{\mathbf{T}_p^T \mathbf{T}_p}{n-1} \right)^{-\frac{1}{2}} \mathbf{p}_{p, k}^T \hat{\mathbf{x}}_{new, k} \right\|^2, \quad (32)$$

其中:  $\mathbf{p}_{p, k}$  是  $\mathbf{P}_p$  的第  $k$  行,  $\hat{\mathbf{x}}_{new, k}$  表示  $\hat{\mathbf{x}}_{new}$  的第  $k$  个元素.

基于  $T_{x_r}^2$  统计量的贡献定义如下<sup>[29]</sup>:

$$C(T_{x_r}^2, k) = \left\| \left( \frac{\mathbf{T}_r^T \mathbf{T}_r}{n-1} \right)^{-\frac{1}{2}} \mathbf{p}_{r, k}^T \tilde{\mathbf{x}}_{new, k} \right\|^2, \quad (33)$$

其中:  $\mathbf{p}_{r, k}$  是  $\mathbf{P}_r$  的第  $k$  行,  $\tilde{\mathbf{x}}_{new, k}$  表示  $\tilde{\mathbf{x}}_{new}$  的第  $k$  个元素.

考虑到复杂系统中历史数据对当前样本的影响, 将贡献重新定义如下:

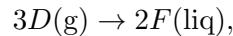
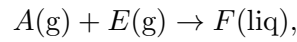
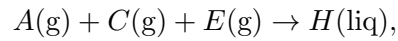
$$C_{Q_{\tilde{x}_z}, k} = \sum_{j=1}^N C(Q_{\tilde{x}_z}, k)_j, \quad (34)$$

$$C_{T_{x_p}^2, k} = \sum_{j=1}^N C(T_{x_p}^2, k)_j, \quad (35)$$

$$C_{T_{x_r}^2, k} = \sum_{j=1}^N C(T_{x_r}^2, k)_j, \quad (36)$$

其中  $N$  表示已获得的样本数, 第  $N$  个样本即为当前样本.

但是, 研究发现, 在无故障时每个过程输入变量对统计量的贡献值并不是均等的. 以田纳西伊士曼过程 (Tennessee Astman process, TEP)<sup>[30]</sup> 为例, TEP 是由伊士曼化学公司为评估过程控制和监测方法提供一个现实的工业过程而创建的, 整个过程由反应器、冷凝器、汽提塔、压缩机和汽液分离器5个单元组成, 包含  $A, B, C, D, E, F, G$  和  $H$  等8种成份, 具体反应如下:



其中:  $g$  表示气态,  $liq$  表示液态,  $F$  为副产物.

如图2所示是利用 MPLS 对 TEP 正常数据进行故障诊断的传统贡献图. 由图可以看出, 正常工况下不同变量对统计量的贡献不同, 最大差异可达8个数量级, 这可能在使用贡献图诊断故障时会产生较大影响, 使得诊断结果不可解释.

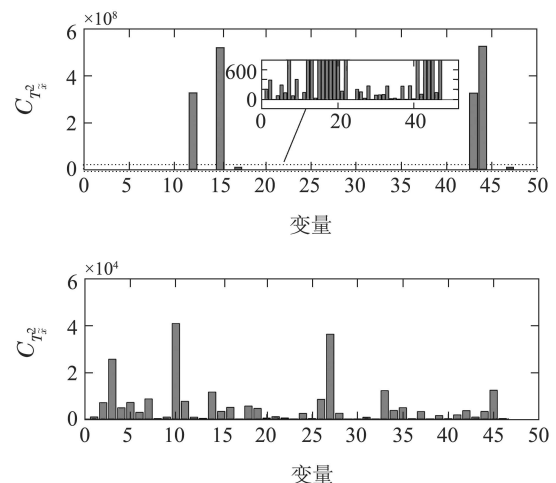


图2 TEP正常数据的MPLS传统贡献图

Fig. 2 Traditional contribution plots of MPLS for normal data of TEP

为了解决上述问题, 本节提出了一种NRC法. 首先由式(31)–(33)在  $n$  个正常训练样本中求出每个样本各变量对相应统计量的贡献值, 然后在所有贡献值中

选取每个变量的最大贡献, 第 $k$ 个变量对各统计量的最大贡献分别记作 $C_{Q_{\bar{x}_z}, \max}^k, C_{T_{x_p}^2, \max}^k, C_{T_{x_r}^2, \max}^k$ . 在得到新样本 $\mathbf{x}_{\text{new}}$ 后, 首先利用式(31)–(33)计算各变量贡献值 $C(Q_{\bar{x}_z}, k), C(T_{x_p}^2, k), C(T_{x_r}^2, k)$ . 然后由 $C(Q_{\bar{x}_z}, k)/C_{Q_{\bar{x}_z}, \max}^k, C(T_{x_p}^2, k)/C_{T_{x_p}^2, \max}^k, C(T_{x_r}^2, k)/C_{T_{x_r}^2, \max}^k$ 进行归一, 归一后的贡献仍记作 $C(Q_{\bar{x}_z}, k), C(T_{x_p}^2, k), C(T_{x_r}^2, k)$ . 最后由式(34)–(36)计算新的贡献值. 图3为MPLS利用NRC对TEP正常数据进行故障诊断结果, 由图可以看出, 经过对贡献归一后, 每个变量对统计量的贡献相对持平, 使得诊断结果误差减小.

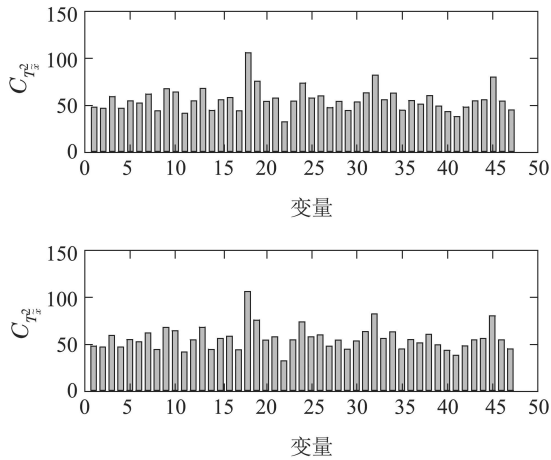


图3 TEP正常数据的MPLS新的相对贡献图  
Fig. 3 New relative contribution plots of MPLS for normal data of TEP

### 4 数值仿真与TEP应用

本节将CMPLS与MPLS<sup>[16]</sup>和EPLS<sup>[17]</sup>方法进行比较, 通过数值案例与TEP验证所提新方法的有效性.

#### 4.1 数值仿真

采用以下模型生成无故障训练样本<sup>[13]</sup>:

$$\begin{cases} \mathbf{x}_k = \mathbf{A}z_k + \mathbf{e}_k, \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k, \end{cases} \quad (37)$$

其中:

$$z_k \sim U([0, 1]), \mathbf{e}_k \sim N(0, 0.2^2), \mathbf{v}_k \sim N(0, 0.01^2),$$

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 4 & 4 & 0 \\ 3 & 0 & 1 & 4 & 1 \\ 1 & 1 & 3 & 0 & 0 \end{bmatrix}^T, \mathbf{C} = \begin{bmatrix} 2 & 2 & 1 & 1 & 0 \\ 3 & 1 & 0 & 4 & 0 \end{bmatrix},$$

$U([0, 1])$ 表示区间 $[0, 1]$ 内的均匀分布.

首先用式(37)生成100个正常的训练样本进行CMPLS建模, 由CPV<sup>[25]</sup>计算可得 $l_p = 2, l_r = 3, l_y = 1$ . 然后在不同的故障情况下, 使用式(38)或式(39)产生100个故障样本用于故障检测. 本小节后续试验中, 前100个样本为正常样本, 后100个样本为故障样本, 使用 $T_{x_p}^2, T_{x_r}^2, Q_{\bar{x}_z}$ 和 $F_y$ 四个统计指标进行过程监

控.

$$\mathbf{x}_k = \mathbf{x}_k^* + \mathbf{\Xi}_x f_x, \quad (38)$$

$$\mathbf{y}_k = \mathbf{y}_k^* + \mathbf{\Xi}_y f_y. \quad (39)$$

式(38)–(39)中,  $\mathbf{x}_k^*$ 和 $\mathbf{y}_k^*$ 分别为无故障输入输出样本,  $\mathbf{\Xi}_x$ 和 $\mathbf{\Xi}_y$ 为故障方向向量,  $f_x$ 和 $f_y$ 为故障幅值. 以下就不同情况的故障进行详细分析.

#### 故障1 发生在IPS中的故障.

为了在IPS中生成故障, 取 $\mathbf{P}_p$ 的第1列为 $\mathbf{\Xi}_x = [0.5269 \ 0.5102 \ 0.4915 \ 0.4611 \ -0.0888]^T, f_x$ 为4. 图4为CMPLS对故障1的检测结果, 可以看出,  $T_{x_p}^2$ 统计量可以有效检测到该故障, 表明在过程输入空间由 $T_{x_p}^2$ 检测到的故障为质量相关故障; 由 $F_y$ 统计量检测结果可以推测该故障还影响了不可预测的质量输出. 由图5所示, 此故障的确影响了质量输出和不可预测的质量输出.

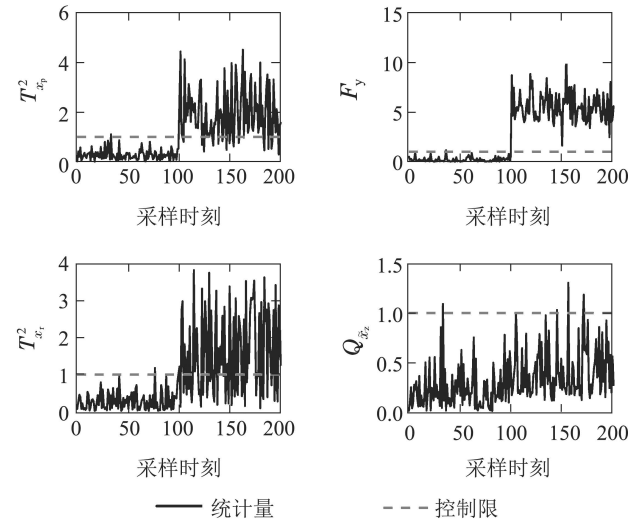


图4 CMPLS对故障1检测结果  
Fig. 4 Detection results of CMPLS for Fault 1

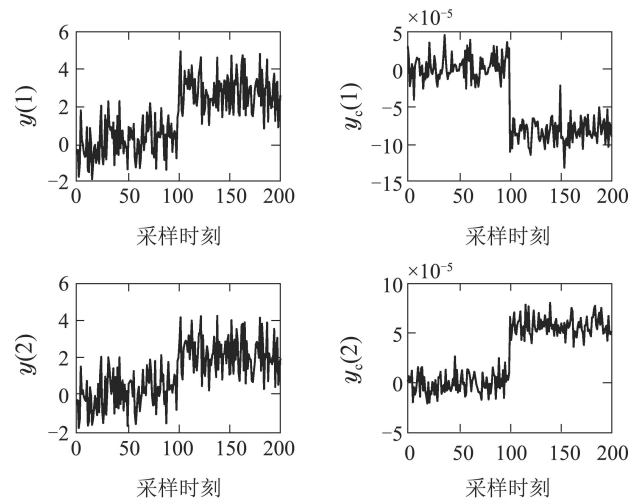


图5 故障1下质量输出与不可预测质量输出  
Fig. 5 Quality output and unpredictable quality output in case of Fault 1

**故障2** 发生在IRPS中的故障。

取  $\Xi_x = [0 \ -0.1497 \ 0.6257 \ -0.5942 \ -0.4827]^T$  为  $P_r$  的第1列,  $f_x$  为3. 图6为CMPLS对故障2的检测结果, 图中  $T_{x_p}^2$  统计量未检测到该故障, 而  $T_{x_r}^2$  统计量检测到该故障, 同时对  $F_y$  统计量有影响, 表明由  $T_{x_r}^2$  统计量检测到的故障为与质量相关的潜在故障, 不直接影响质量输出, 但对无法预测的质量输出有一定影响. 由图7所示, 该故障确实不影响质量输出, 但会影响不可预测的质量输出.

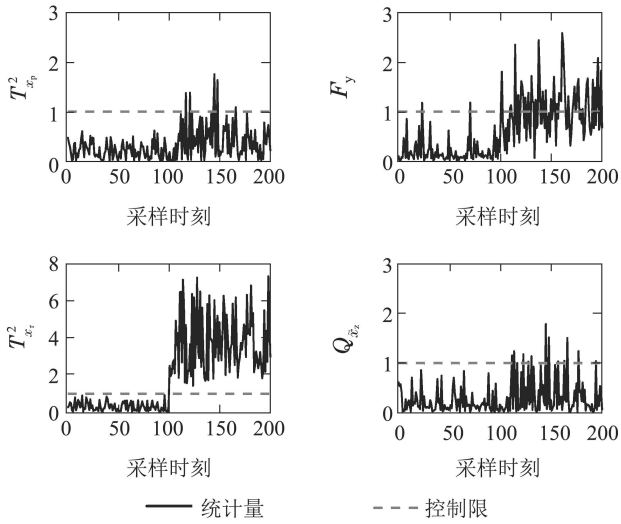


图6 CMPLS对故障2检测结果

Fig. 6 Detection results of CMPLS for Fault 2

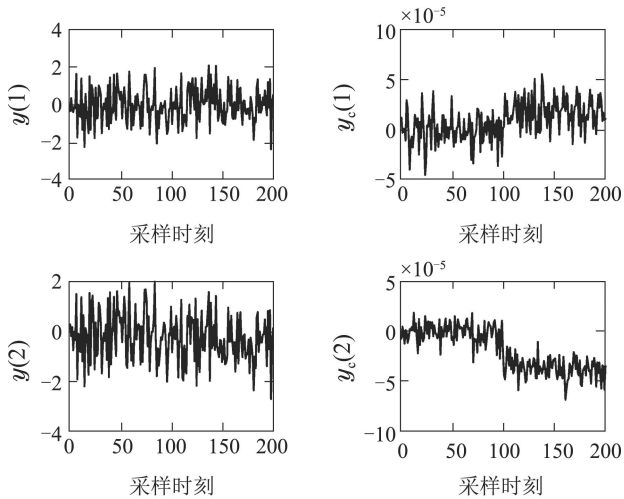


图7 故障2下质量输出与不可预测质量输出

Fig. 7 Quality output and unpredictable quality output in case of Fault 2

**故障3** 发生在IRS中的故障。

如式(19)所示, IRS的基向量为  $\hat{\Gamma}_\varphi \hat{\Gamma}_\varphi^T (I - P_p P_p^T) + (I - \hat{\Gamma}_\varphi \hat{\Gamma}_\varphi^T) (I - P_r P_r^T)$  的非零奇异值对应的左奇异向量, 因此可以取IRS的一个基向量为故障方向  $\Xi_x = [0.5288 \ -0.4903 \ -0.2517 \ -0.5201 \ 0.3823]^T$ ,

$f_x$  为3. CMPLS对故障3的检测结果如图8所示, 只有  $Q_{x_z}$  统计量检测到故障, 则说明该故障与质量输出无关, 但与过程输入相关.

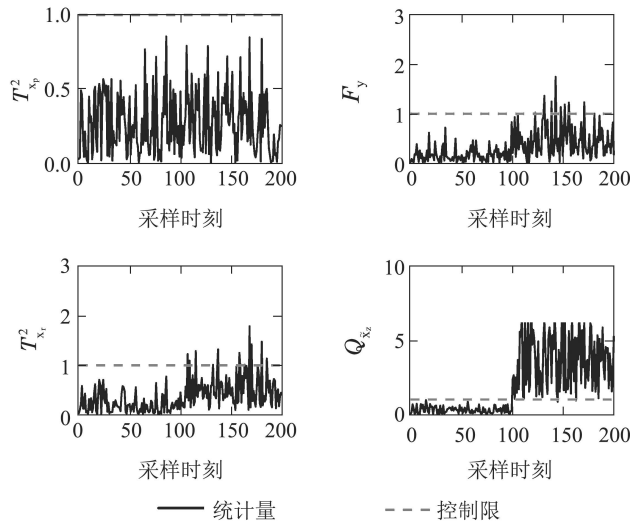


图8 CMPLS对故障3检测结果

Fig. 8 Detection results of CMPLS for Fault 3

**故障4** 发生在OPS中的故障。

取  $\Xi_y = [-0.7071 \ 0.7071]^T$  为  $P_y$  的第1列,  $f_y$  为0.5. CMPLS对故障4的检测结果如图9所示. 该故障发生在不可预测的质量输出空间中, 与过程输入无关, 图9中只有  $F_y$  统计量检测到故障, 这意味着当其他统计量均未监测到故障时,  $F_y$  统计量监测到的故障为与过程输入无关, 但与不可预测的质量输出相关的故障. 图10与图11分别为MPLS与EPLS对故障4的检测结果, 由图可以看出, 基于MPLS与EPLS的故障检测方法无法检测到此类故障.

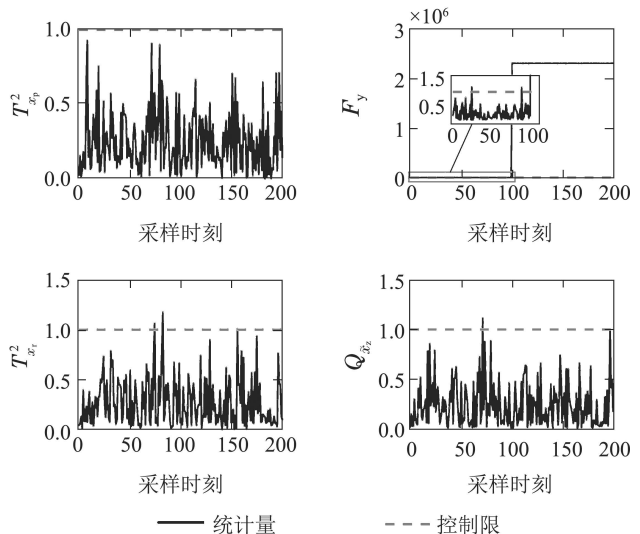


图9 CMPLS对故障4检测结果

Fig. 9 Detection results of CMPLS for Fault 4

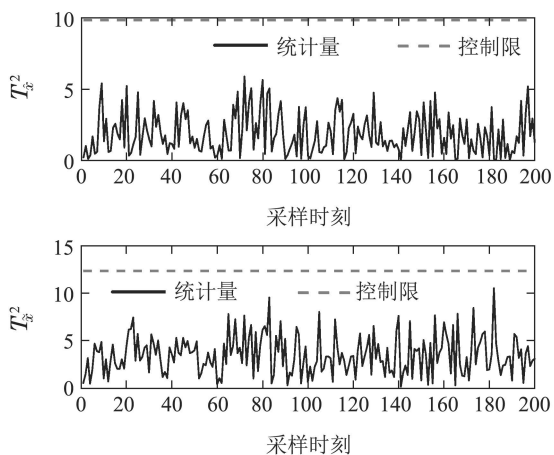


图 10 MPLS对故障4检测结果

Fig. 10 Detection results of MPLS for Fault 4

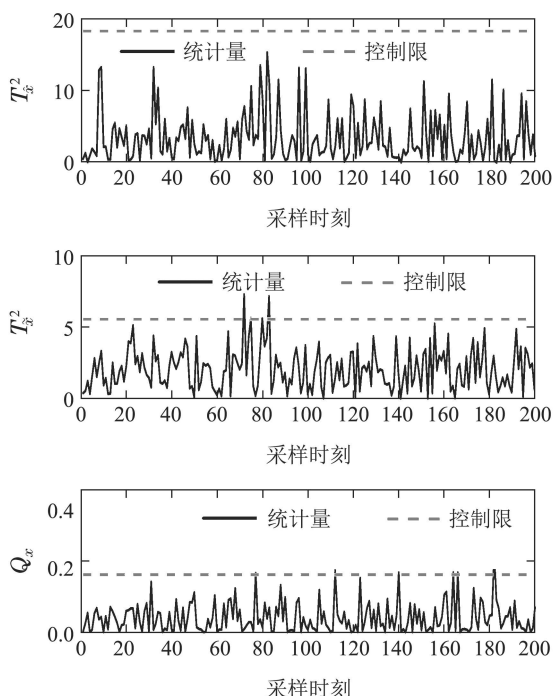


图 11 EPLS对故障4检测结果

Fig. 11 Detection results of EPLS for Fault 4

### 4.2 TEP应用

本小节利用田纳西伊士曼过程(TEP)<sup>[30]</sup>验证所提CMPLS方法的有效性. 该过程包含41个测量变量(measurement variable, MEAS)和12个操纵变量(manipulated variable, MV), 测量变量包含22个连续变量和19个成分变量. 该过程共有21种故障, 定义为IDV(1)–IDV(21), 其中16个已知故障, 5个未知故障. 本节实验中, 过程输入变量选择36个MEAS(1–36)(对应实验中变量1–36)和11个MV(1–11)(对应实验中变量37–47), 质量输出变量选择5个MEAS(37–41). 后续实验中, 前500个样本为正常数据, 后480个样本为含有故障的数据. 通过CPV<sup>[25]</sup>计算可得CMPLS中 $l_p = 2$ ,  $l_r = 24$ ,  $l_y = 4$ .

### 4.2.1 故障检测

在本小节以故障IDV(2)与IDV(4)为例利用MPLS与CMPLS分别进行故障检测, 以验证算法有效性. 由文献[16], IDV(2)为质量相关故障, IDV(4)为质量无关故障. IDV(2)为汽提塔入口处B组分含量发生阶跃变化, 图12与图13分别是MPLS与CMPLS对IDV(2)的过程监测结果, MPLS的两个检测指标和CMPLS的4个检测指标均检测到故障.

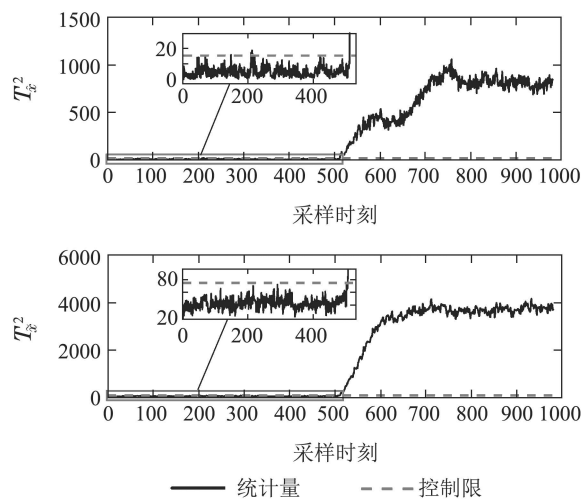


图 12 MPLS对IDV(2)过程监测结果

Fig. 12 Process monitoring results of MPLS for IDV(2)

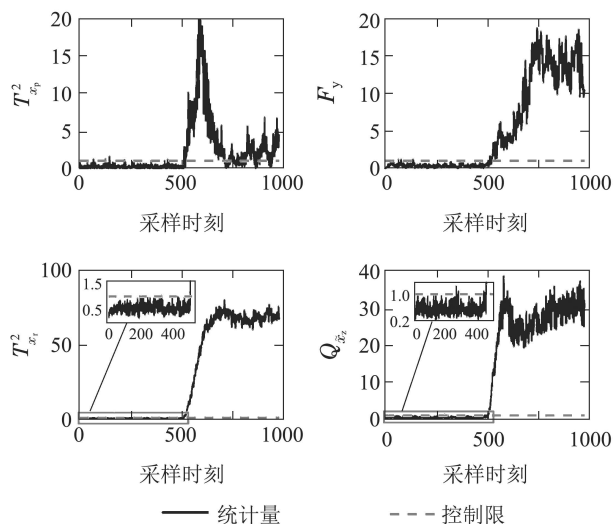


图 13 CMPLS对IDV(2)过程监测结果

Fig. 13 Process monitoring results of CMPLS for IDV(2)

对于CMPLS, 当IDV(2)发生后, 与质量相关的 $T_{x_p}^2$ 统计量在一段时间后向正常值下降, 这表明故障对质量输出变量的影响趋于回归正常, 这是由于过程中通过反馈控制器调节减少故障对质量输出的影响. 值得注意的是, 当 $T_{x_p}^2$ 统计量在故障阶段降低时, 其他统计量却在上升或一直处于较高水平, 这是因为反馈控制器调节虽然降低了故障对质量输出的影响, 但故障并没有消除, 会对其他统计量产生较大影响. 在检测故障阶段, 应该根据第2.1节检测策略进行综合判断,



当 $T_{x_p}^2$  统计量降到控制限以下时,可以看出此时故障IDV(2)不会直接影响质量输出,但会影响过程输入以及不可预测的质量输出.在MPLS的监测结果中无法观察到反馈对质量变化的影响.

IDV(4)是反应器冷却水入口温度发生了阶跃变化,这种扰动不会影响产品质量,图14与图15分别是MPLS与CMPLS对IDV(4)的过程监测结果.由图可以看出,CMPLS中 $T_{x_p}^2 < J_{th,T_{x_p}^2}$ 且 $Q_{\bar{x}_z} > J_{th,Q_{\bar{x}_z}}$ ,根据检测策略,该故障为与质量输出无关,但与过程输入相关的故障;同时 $T_{x_r}^2$ 有一定的报警,说明该故障不直接影响质量输出,这一点体现在 $T_{x_p}^2$  统计量中,但对不可预测的质量输出有一定影响,因此 $F_y$  统计量在故障阶段有一定波动.在MPLS的监测结果中只能判断IDV(4)为质量无关故障,得不到更多信息.

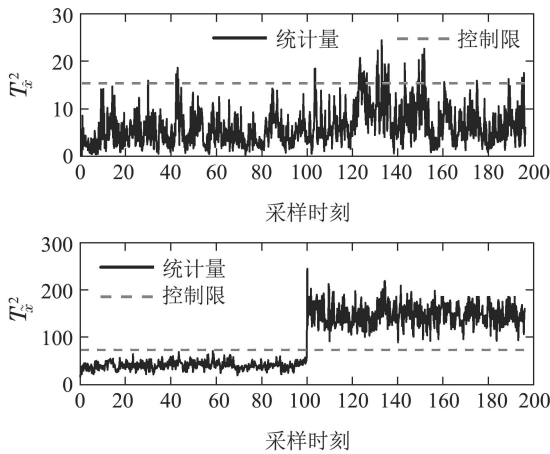


图 14 MPLS对IDV(4)过程监测结果

Fig. 14 Process monitoring results of MPLS for IDV(4)

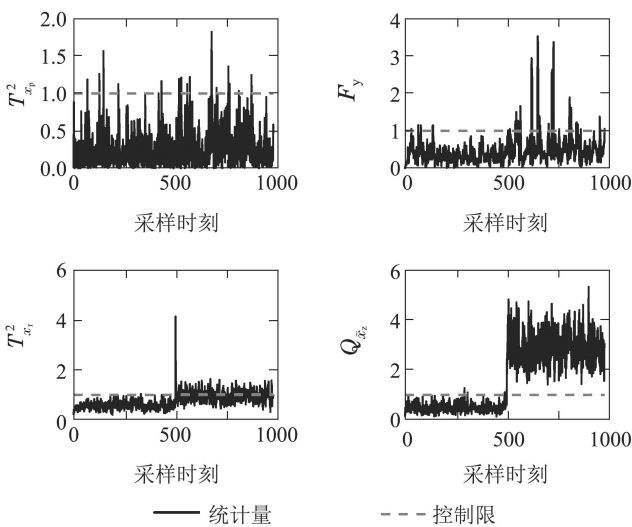


图 15 CMPLS对IDV(4)过程监测结果

Fig. 15 Process monitoring results of CMPLS for IDV(4)

4.2.2 故障诊断

在进行故障诊断之前,需要检测到故障.本小节以故障IDV(14)为例,首先利用MPLS与CMPLS对

IDV(14)进行过程监测,然后分别通过MPLS结合传统贡献图,MPLS结合NRC以及CMPLS结合NRC对IDV(14)进行故障诊断,验证所提新方法的有效性.

在TEP中,IDV(14)与反应器冷却水阀门有关.当IDV(14)发生时,反应器冷却水阀门将卡住保持不动,这将使得过程测量变量中反应器温度(MEAS(9)),反应器冷却水出口温度(MEAS(21))以及过程操纵变量中反应器冷水流量(MV(10))受到很大影响<sup>[31]</sup>,分别对应实验中变量9,21,46.图16和图17是MPLS与CMPLS对IDV(14)的过程监测结果.由文献[16]已知,IDV(14)不会对质量产生影响,而图16中,MPLS的质量相关检测指标有较高的误报.由图17可得,CMPLS中 $T_{x_p}^2 < J_{th,T_{x_p}^2}$ 且 $Q_{\bar{x}_z} > J_{th,Q_{\bar{x}_z}}$ ,根据检测策略,该故障为与过程输入相关的质量无关故障;同时 $T_{x_r}^2 > J_{th,T_{x_r}^2}$ ,说明该故障对不可预测的质量输出有一定影响,因此 $F_y$  统计量在故障阶段有一定报警.CMPLS对IDV(14)给出了比MPLS更加准确而又全面的检测结果.

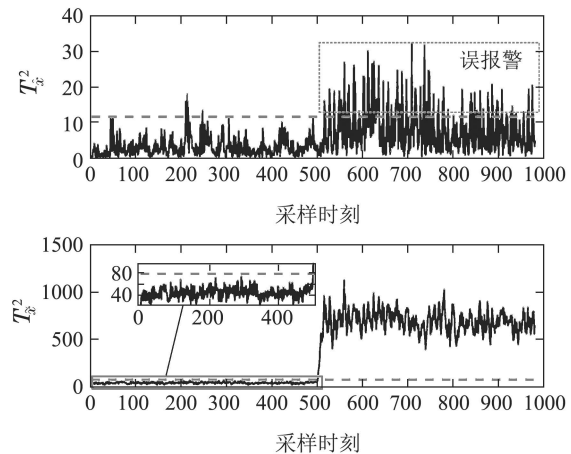


图 16 MPLS对IDV(14)过程监测结果

Fig. 16 Process monitoring results of MPLS for IDV(14)

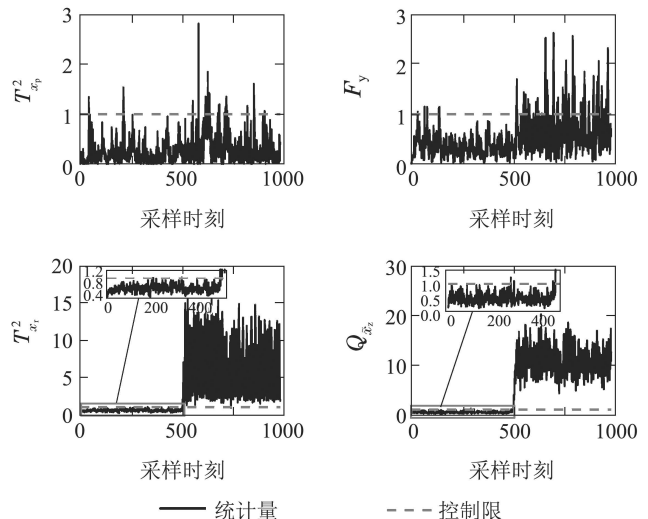


图 17 CMPLS对IDV(14)过程监测结果

Fig. 17 Process monitoring results of CMPLS for IDV(14)

图18是 MPLS 结合传统贡献图对IDV(14)的诊断结果. 由图可以看出, 变量12, 15, 43与44对质量相关统计量的贡献极大, 但IDV(14)主要影响的并不是这几个变量. 变量9与21对质量无关统计量的贡献最大, 这在一定程度上可以解释故障引起的变化, 但并不全面和准确. 除了变量9和21外, 其他变量也有较大贡献, 同时, 变量46的贡献却很低, 这一结果很难解释. 实际上, 结合图2可以看出, 对质量相关统计量的贡献就是由正常情况下的各变量对统计量的不均等造成的, 对质量无关统计量也有同样的影响.

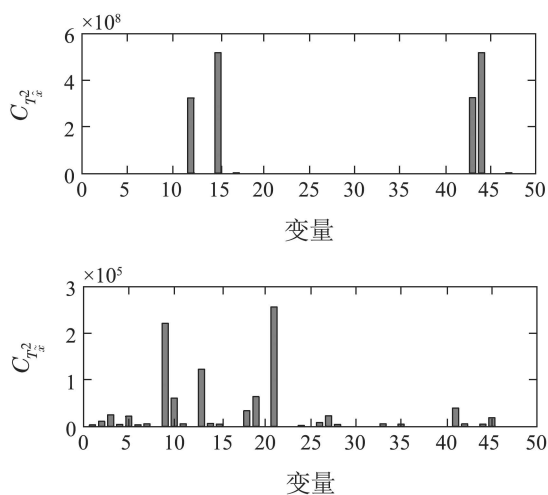


图 18 IDV(14)的MPLS传统贡献图

Fig. 18 Traditional contribution plots of MPLS for IDV(14)

图19与图20分别为MPLS 与 CMPLS 结合NRC对IDV(14)的故障诊断结果. 由图19, MPLS结合NRC在两个统计量中均准确的定位了故障变量, 相比较MPLS结合传统贡献图, 消除了传统贡献图中各变量正常贡献不均的影响. 由图20, CMPLS结合NRC诊断结果中, 各变量对 $T_{x_p}^2$  统计量贡献均等, 这是由于IDV(14)为质量输出无关故障, 在IPS中无故障信息; 而 $T_{x_r}^2$  和 $Q_{x_z}$  统计量能够监测到该故障, 同时在IRPS与IRS中准确定位到相应故障变量, 在IRPS中主要受影响的变量为9, 21, 46, 在IRS中主要受影响的变量为21. 比较图19与图20可以得到, 利用CMPLS进行故障诊断可以区分出故障是否质量输出相关, 而利用MPLS进行故障诊断无法确定故障是否质量输出相关. 综合来看, 所提CMPLS与NRC均比MPLS与传统贡献图的性能更好.

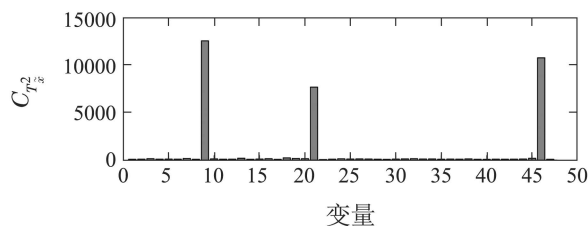
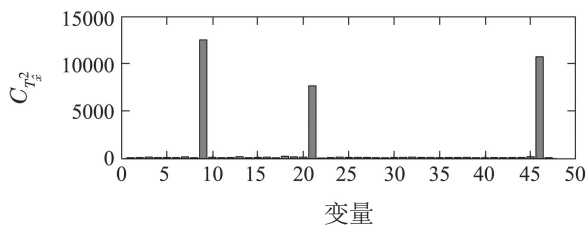


图 19 IDV(14)的MPLS新的相对贡献图

Fig. 19 New relative contribution plots of MPLS for IDV(14)

图 20 IDV(14)的CMPLS新的相对贡献图

Fig. 20 New relative contribution plots of CMPLS for IDV(14)

### 5 结论

本文提出了一种新的CMPLS算法来检测质量输出相关和过程输入相关的故障, 并利用所提的新的相对贡献图法成功诊断出故障变量. 通过CMPLS同时将输入输出数据映射到与过程相关或质量相关的多个子空间, 针对每个子空间分别提出了对应的故障检测指标, 用于各类故障的检测报警. 该方法对可预测及不可预测的质量输出子空间中发生的故障以及影响过程输入空间并可能影响质量输出的潜在故障进行了全面监测, 同时避免了大量的迭代过程. 重新定义了一种新的相对贡献图, 消除了正常工况下各变量向检测指标贡献不均对诊断结果所造成的不良影响. 通过数值仿真和TEP验证了所提方法的优异性能.

### 参考文献:

[1] LANNOY A, PROCACCIA H. Expertise, safety, reliability, and decision making: practical industrial experience. *Environment Systems and Decisions*, 2014, 34(2): 259 – 276.  
 [2] CHU Xiaoyan, NIAN Xiaohong, LIU Jingjing. Fault detection and isolation for multi-motor web-winding system via robust sliding mode observers. *Control Theory & Applications*, 2018, 35(6): 795 – 804.

- (楚晓艳, 年晓红, 刘静静. 基于鲁棒滑模观测器的多电机卷绕系统故障检测和隔离. *控制理论与应用*, 2018, 35(6): 795 – 804.)
- [3] YIN S, WANG G, GAO H. Data-driven process monitoring based on modified orthogonal projections to latent structures. *IEEE Transactions on Control Systems Technology*, 2016, 24(4): 1480 – 1487.
- [4] FENG Liwei, LI Yuan, ZHANG Cheng, et al. Time-space neighborhood standardization-local outlier factor based fault detection for complex process. *Control Theory & Applications*, 2020, 37(3): 651 – 657.  
(冯立伟, 李元, 张成, 等. 基于时空近邻标准化和局部离群因子的复杂过程故障检测. *控制理论与应用*, 2020, 37(3): 651 – 657.)
- [5] LUO H, ZHAO H, YIN S. Data-driven design of fog-computing-aided process monitoring system for large-scale industrial processes. *IEEE Transactions on Industrial Informatics*, 2018, 14(10): 4631 – 4641.
- [6] HE R, CHEN G, DONG C, et al. Data-driven digital twin technology for optimized control in process systems. *ISA Transactions*, 2019, 95: 221 – 234.
- [7] GE Z, SONG Z, GAO F. Review of recent research on data-based process monitoring. *Industrial & Engineering Chemistry Research*, 2013, 52(10): 3543 – 3562.
- [8] JOLLIFFE I T. *Principal Component Analysis and Factor Analysis*. New York: Springer, 1986.
- [9] WOLD S, MARTENS H, WOLD H. The multivariate calibration problem in chemistry solved by the PLS method. *Lecture Notes in Mathematics*, 1983, 973: 286 – 293.
- [10] WISE B M, GALLAGHER N B. The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*, 1996, 6(6): 329 – 348.
- [11] QIN S J. Statistical process monitoring: basics and beyond. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 2003, 17(8): 480 – 502.
- [12] WANG L, MA C, DAI S. A distributed multichannel signal acquisition system based on DSP. *Proceedings of the 11th International Conference on Computer and Automation Engineering*. New York: Association for Computing Machinery, 2019: 109 – 113.
- [13] QIN S J, ZHENG Y. Quality-relevant and process-relevant fault monitoring with concurrent projection to latent structures. *AIChE Journal*, 2013, 59(2): 496 – 504.
- [14] YIN S, DING S X, XIE X, et al. A review on basic data-driven approaches for industrial process monitoring. *IEEE Transactions on Industrial Electronics*, 2014, 61(11): 6418 – 6428.
- [15] ZHOU D H, LI G, QIN S J. Total projection to latent structures for process monitoring. *AIChE Journal*, 2010, 56(1): 168 – 178.
- [16] YIN S, DING S X, ZHANG P, et al. Study on modifications of PLS approach for process monitoring. *IFAC Proceedings Volumes*, 2011, 44(1): 12389 – 12394.
- [17] ZHANG K, PENG K, YOU B, et al. Quality-relevant fault monitoring based on efficient projection to latent structures with application to hot strip mill process. *IET Control Theory & Applications*, 2015, 9(7): 1135 – 1145.
- [18] HE B, CHEN T, YANG X. Root cause analysis in multivariate statistical process monitoring: Integrating reconstruction-based multivariate contribution analysis with fuzzy-signed directed graphs. *Computers & Chemical Engineering*, 2014, 64: 167 – 177.
- [19] ZHOU Z, WEN C, YANG C. Fault isolation based on  $k$ -nearest neighbor rule for industrial processes. *IEEE Transactions on Industrial Electronics*, 2016, 63(4): 2578 – 2586.
- [20] DENG Jiawei, DENG Xiaogang, CAO Yuping, et al. Incipient fault diagnosis method of nonlinear chemical process based on weighted statistical local KPCA. *CIESC Journal*, 2019, 70(7): 2594 – 2605.  
(邓佳伟, 邓晓刚, 曹玉苹, 等. 基于加权统计局部核主元分析的非线性化工过程微小故障诊断方法. *化工学报*, 2019, 70(7): 2594 – 2605.)
- [21] WANG Jinrui, XIE Lirong, WANG Zhongqiang, et al. Fault diagnosis of rolling bearing based on MEEMD–DHENN. *Journal of Mechanical Transmission*, 2018, 42(3): 139 – 143.  
(王晋瑞, 谢丽蓉, 王忠强, 等. 基于MEEMD–DHENN的滚动轴承故障诊断. *机械传动*, 2018, 42(3): 139 – 143.)
- [22] MACGREGOR J F, KOURTI T. Statistical process control of multivariate processes. *Control Engineering Practice*, 1995, 3(3): 403 – 414.
- [23] LI G, QIN S Z, JI Y D, et al. Total PLS based contribution plots for fault diagnosis. *Acta Automatica Sinica*, 2009, 35(6): 759 – 765.
- [24] WANG G, YIN S. Quality-related fault detection approach based on orthogonal signal correction and modified PLS. *IEEE Transactions on Industrial Informatics*, 2017, 11(2): 398 – 405.
- [25] ZHOU Donghua, LI Gang, LI Yuan. *Data-Driven Fault Diagnosis Technology for Industrial Process*. Beijing: Science Press, 2011.  
(周东华, 李钢, 李元. 数据驱动的工业过程故障诊断技术. 北京: 科学出版社, 2011.)
- [26] MUDHOLKAR J G S. Control procedures for residuals associated with principal component analysis. *Technometrics*, 1979, 21(3): 341 – 349.
- [27] VIDAL-PUIG S, VITALE R, FERRER A. Data-driven supervised fault diagnosis methods based on latent variable models: a comparative study. *Chemometrics and Intelligent Laboratory Systems*, 2019, 187: 41 – 52.
- [28] TAN R M, CAO Y. Multi-layer contribution propagation analysis for fault diagnosis. *International Journal of Automation and Computing*, 2019, 16(1): 40 – 51.
- [29] QIN S J, VALLE S, PIOVOSO M J. On unifying multiblock analysis with application to decentralized process monitoring. *Journal of Chemometrics*, 2001, 15(9): 715 – 742.
- [30] DOWNS J J, VOGEL E F. A plant-wide industrial process control problem. *Computers & Chemical Engineering*, 1993, 17(3): 245 – 255.
- [31] LI J, ZHANG L, JIAO J, et al. Quality-related fault diagnosis based on improved PLS for industrial process. *The 34th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. Jinzhou, China: IEEE, 2019: 296 – 301.

#### 作者简介:

**李强** 硕士研究生, 目前研究方向为系统特征提取、多元统计过程监控, E-mail: leeqang@yeah.net;

**孔祥玉** 教授, 博士生导师, 目前研究方向为控制系统故障诊断、非线性系统辨识、复杂系统故障预报与寿命预测等, E-mail: xianguyukong01@163.com;

**罗家宇** 博士研究生, 目前研究方向为数据驱动的复杂系统下故障监测与诊断等, E-mail: 540629964@qq.com;

**解建** 硕士研究生, 目前研究方向为系统特征提取、多元统计过程监控, E-mail: hdxiejian@163.com.